



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2013

Towards an accurate mass function for precision cosmology

Reed, D S ; Smith, R E ; Potter, D ; Schneider, A ; Stadel, J ; Moore, B

Abstract: Cosmological surveys aim to use the evolution of the abundance of galaxy clusters to accurately constrain the cosmological model. In the context of Λ cold dark matter, we show that it is possible to achieve the required per cent level accuracy in the halo mass function with gravity-only cosmological simulations, and we provide simulation start and run parameter guidelines for doing so. Some previous works have had sufficient statistical precision, but lacked robust verification of absolute accuracy. Convergence tests of the mass function with, for example, simulation start redshift can exhibit false convergence of the mass function due to counteracting errors, potentially misleading one to infer overly optimistic estimations of simulation accuracy. Per cent level accuracy is possible if initial condition particle mapping uses second-order Lagrangian perturbation theory, and if the start epoch is between 10 and 50 expansion factors before the epoch of halo formation of interest. The mass function for haloes with fewer than ~ 1000 particles is highly sensitive to simulation parameters and start redshift, implying a practical minimum mass resolution limit due to mass discreteness. The narrow range in converged start redshift suggests that it is not presently possible for a single simulation to capture accurately the cluster mass function while also starting early enough to model accurately the numbers of reionization era galaxies, whose baryon feedback processes may affect later cluster properties. Ultimately, to fully exploit current and future cosmological surveys will require accurate modelling of baryon physics and observable properties, a formidable challenge for which accurate gravity-only simulations are just an initial step.

DOI: <https://doi.org/10.1093/mnras/stt301>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-80943>

Journal Article

Published Version

Originally published at:

Reed, D S; Smith, R E; Potter, D; Schneider, A; Stadel, J; Moore, B (2013). Towards an accurate mass function for precision cosmology. *Monthly Notices of the Royal Astronomical Society*, 431(2):1866-1882.

DOI: <https://doi.org/10.1093/mnras/stt301>

Towards an accurate mass function for precision cosmology

Darren S. Reed,¹★ Robert E. Smith,^{1,2} Doug Potter,¹ Aurel Schneider,¹ Joachim Stadel¹ and Ben Moore¹

¹*Institute for Theoretical Physics, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*

²*Argelander-Institute for Astronomy, Auf dem Hügel 71, D-53121 Bonn, Germany*

Accepted 2013 February 16. Received 2013 January 22; in original form 2012 June 22

ABSTRACT

Cosmological surveys aim to use the evolution of the abundance of galaxy clusters to accurately constrain the cosmological model. In the context of Λ cold dark matter, we show that it is possible to achieve the required per cent level accuracy in the halo mass function with gravity-only cosmological simulations, and we provide simulation start and run parameter guidelines for doing so. Some previous works have had sufficient statistical precision, but lacked robust verification of absolute accuracy. Convergence tests of the mass function with, for example, simulation start redshift can exhibit false convergence of the mass function due to counteracting errors, potentially misleading one to infer overly optimistic estimations of simulation accuracy. Per cent level accuracy is possible if initial condition particle mapping uses second-order Lagrangian perturbation theory, and if the start epoch is between 10 and 50 expansion factors before the epoch of halo formation of interest. The mass function for haloes with fewer than ~ 1000 particles is highly sensitive to simulation parameters and start redshift, implying a practical minimum mass resolution limit due to mass discreteness. The narrow range in converged start redshift suggests that it is not presently possible for a single simulation to capture accurately the cluster mass function while also starting early enough to model accurately the numbers of reionization era galaxies, whose baryon feedback processes may affect later cluster properties. Ultimately, to fully exploit current and future cosmological surveys will require accurate modelling of baryon physics and observable properties, a formidable challenge for which accurate gravity-only simulations are just an initial step.

Key words: galaxies: haloes – cosmology: theory – dark matter.

1 INTRODUCTION

In the vacuum energy-dominated cold dark matter (Λ CDM) cosmological model (Komatsu et al. 2011), large-scale structures form through the amplification of small density fluctuations via gravitational instability. At early times, this amplification can be followed using linear perturbation theory of the general relativistic equations of motion for the field. At late times, owing to the non-linearities in the equations, and after shell-crossing, the dynamics may only be accurately followed using numerical simulations. Overdense regions of the density field, whose dynamics have broken away from the evolution of the background space–time and have reached some state of virial equilibrium, are commonly referred to as dark matter haloes.

The growth rate of large-scale structures is directly sensitive to the expansion rate of the Universe, and hence the cosmological parameters. One can show theoretically, through the excursion set

formalism (Press & Schechter 1974; Bond et al. 1991; Sheth & Tormen 1999), that the number of haloes is also sensitive to cosmological parameters, and importantly for future surveys, the presence of ‘dark energy’ (Wang & Steinhardt 1998; Haiman, Mohr & Holder 2001; Lima & Hu 2004; Marian & Bernstein 2006; Cunha, Huterer & Doré 2010; Courtin et al. 2011). This forecast cosmological sensitivity has recently been verified through direct testing with N -body simulations (Smith & Marian 2011).

The amount of cosmological information that can be extracted from cluster number counts is limited by our ability to detect signal-to-noise ratio peaks in our observational survey – i.e. associate galaxies to groups, identify groups relative to an X-ray background noise level, etc. The lower that one can push the minimum detectable mass, the more cosmological information can be extracted from the survey. This comes under the proviso that one can accurately calibrate the true–observable mass relationship (Lima & Hu 2005; Marian, Smith & Bernstein 2009; Rozo et al. 2009; Mandelbaum et al. 2010; Oguri & Takada 2011; Angulo et al. 2012). The numbers of rare haloes are also sensitive to the level of non-Gaussianity in the primordial density field due to its effect upon the tail of

★ E-mail: reed@physik.uzh.ch

extreme density fluctuations (Matarrese, Verde & Jimenez 2000; Marian et al. 2011). Cluster counts are also sensitive to the total neutrino mass (e.g. Wang et al. 2005; Carbone et al. 2012; Shimon et al. 2012). Thus, surveys that promise to accurately measure the evolution of the abundance of groups and clusters also have the potential to help probe fundamental physics. Accurate theoretical predictions for the cluster mass function and its dependence on cosmology are therefore essential to fully exploit next-generation cluster surveys.

Current cosmological constraints from clusters come from: Vikhlinin et al. (2009), Vanderlinde et al. (2010), Rozo et al. (2010), Sehgal et al. (2011), Allen, Evrard & Mantz (2011), Planck Collaboration (2011). Over the next decade there will be a number of large surveys that will aim to strongly constrain the cosmological model through the abundance of clusters: in the X-ray there will be eROSITA (Pillepich, Porciani & Reiprich 2012), with the Sunyaev–Zel’dovich method there will be *Planck*,¹ in the optical using the weak lensing method there will be DES,² Euclid (Laureijs et al. 2011) and LSST.³ Several authors have estimated the requirements on the theoretical accuracy of the halo mass function to achieve the statistically limited constraints on cosmological parameters. Wu, Zentner & Wechsler (2010) point out that, in order to constrain time evolving dark energy models for DES, the theoretical mass function must be known with an accuracy $\lesssim 0.5$ per cent.

In this paper we address the question: what are the correct numerical parameters needed to achieve per cent level accuracy in the mass function in a cosmological simulation? Large simulation volumes (whether by a single simulation or by multiple realizations) are able to reduce statistical uncertainties due to finite halo numbers. However, large absolute volumes are needed to reduce systematic and statistical errors associated with poor sampling of large-scale density modes (e.g. Barkana & Loeb 2004; Bagla & Ray 2005; Power & Kneib 2006; Reed et al. 2007; Crocce et al. 2010; Smith & Marian 2011). Over the past decade, impressive statistical precision in the halo mass function has been achieved using suites of cosmological simulations (Jenkins et al. 2001; Warren et al. 2006; Reed et al. 2007; Tinker et al. 2008; Crocce et al. 2010; Bhattacharya et al. 2011; Smith & Marian 2011; Angulo et al. 2012; Alimi et al. 2012; Iliev et al. 2012; Watson et al. 2012, and others). However, statistical precision does not imply accuracy, even when considering gravity-only simulations. Sources of systematic error include finite simulation volume, force resolution, mass resolution and discreteness effects, time-stepping, halo finding, initial condition particle mapping and start redshift. Recent progress includes Crocce et al. (2010) and Bhattacharya et al. (2011), who each address many of the systematic uncertainties and determine a halo mass function with an estimated accuracy of ~ 2 per cent from a suite of large gravity-only cosmological boxes, though their results differ by significantly more than this for haloes larger than $\sim 10^{15} h^{-1} M_{\odot}$. Moreover, neither approach has taken into account the full covariance matrix of mass function estimates when deriving best-fitting parameters (Smith & Marian 2011).

As a first step on the path towards producing an accurate mass function in the observational plane, we limit ourselves to demonstrating how per cent level accuracy in gravity-only (i.e. collisionless) simulations (wherein baryons are present but interact only via gravity) can be accomplished. We show how to set up initial con-

ditions so that per cent level accuracy can be achieved. We also isolate and test the run parameters that control force resolution (force-softening and tree-opening angle) and time-step size, allowing us to determine the required values to achieve per cent level convergence. Finally, we ask: how many particles do we need to sample the halo mass distribution, in order to obtain a mass function accurate to better than $\lesssim 1$ per cent.

The paper breaks down as follows: in Section 2 we discuss setting up the initial conditions for the structure formation simulations and the parameters used to run the N -body codes. In Section 3 we describe the suite of N -body simulations performed and halo identification. In Section 4 we present the results for the halo mass function and its convergence with simulation parameters. In Section 5 we explore the variation of the halo mass function with the method for generating the initial conditions. We also make a comparison between the results obtained from two well-known N -body codes. In Section 6 we explore the convergence of the matter power spectrum and the one-point probability density function (PDF) of matter fluctuations. In Section 7 we discuss the remaining challenges of obtaining better than 1 per cent accurate mass functions from structure formation simulations. In Section 8 we summarize our findings and draw up a set of guidelines for obtaining accurate gravity-only mass functions.

2 SIMULATING STRUCTURE FORMATION

2.1 Initial conditions

In order to set up a simulation, we must first select the cosmological model and the probability distribution of the primordial density perturbations. In this study we shall work within the context of the Λ CDM paradigm and assume that the initial density modes are described by a Gaussian random field. The statistics of the field are thus fully specified by the power spectrum. Hence, a particular realization of the density field in Fourier space may be obtained by drawing a set of uniform random phases and assigning amplitudes drawn from the Rayleigh distribution (Efstathiou et al. 1985), or through the convolution of white noise with a filter that is related to the power spectrum (Bertschinger 2001).

A given density field must then be converted into a particle distribution, and several techniques for doing this have been discussed in the literature (e.g. Efstathiou et al. 1985; Scoccimarro 1998; Bertschinger 1999, 2001; Crocce, Puellos & Scoccimarro 2006). The traditional approach is to place particles on a uniform lattice, and these are then displaced off the initial points \mathbf{q} using a displacement field $\Psi(\mathbf{q})$ that encodes all of the statistical properties of the density field. Hence, initial (Lagrangian) and final (Eulerian) positions, \mathbf{x} , are related through:

$$\mathbf{x} = \mathbf{q} + \Psi(\mathbf{q}, \tau), \quad (1)$$

where the coordinates \mathbf{x} are a solution to the equation of motion:

$$\frac{d^2 \mathbf{x}}{d\tau^2} + \mathcal{H}(\tau) \frac{d\mathbf{x}}{d\tau} = -\nabla \Phi \quad (2)$$

where in the above $d\tau = dt/a(t)$ is conformal time, $\mathcal{H} = aH(a)$ and Φ is the peculiar gravitational potential. The solution for Ψ is perturbative, and each order can be found through iteration with solutions of lower order. In terms of the initial density field, and up to second order, the solutions may be written (Zel’dovich 1970; Buchert 1994; Buchert, Melott & Weiss 1994; Bouchet et al. 1995; Scoccimarro 1998):

$$\Psi(\mathbf{q}, \tau) = -D_1(a) \nabla_{\mathbf{q}} \phi^{(1)}(\mathbf{q}) + D_2(a) \nabla_{\mathbf{q}} \phi^{(2)}(\mathbf{q}), \quad (3)$$

¹ www.rssd.esa.int/planck

² www.darkenergysurvey.org

³ www.lsst.org/lsst

where $D_1(a)$ and $D_2(a) \approx -3D_1^2(a)/7$ are the first- and second-order growth factors suitable for Λ CDM. The potentials $\phi^{(1)}(\mathbf{q})$ and $\phi^{(2)}(\mathbf{q})$ can be found through iteratively solving the Poisson equations:

$$\nabla_q^2 \phi^{(1)}(\mathbf{q}) = \delta^{(1)}(\mathbf{q}); \quad (4)$$

$$\nabla_q^2 \phi^{(2)}(\mathbf{q}) = \sum_{i>j}^3 \left\{ \phi_{,ii}^{(1)}(\mathbf{q}) \phi_{,jj}^{(1)}(\mathbf{q}) - \left[\phi_{,ij}^{(1)} \right]^2 \right\}, \quad (5)$$

where $\phi_{,ij} \equiv \partial^2 \phi / \partial q_i \partial q_j$. The linear solutions for Ψ , with $\phi^{(2)} = 0$, yield the traditional Zel'Dovich approximation, which we refer to as 1LPT, and the second order solutions, with $\phi^{(2)}$, we refer to as 2LPT. Scoccimarro (1998) gave a detailed prescription for implementing 2LPT displacements in simulations, and we make use of a slightly modified version of the publicly available code 2LPT.⁴ Crocce et al. (2006) demonstrated that 2LPT reduces numerical 'transients' to the level where an accurate representation of the halo mass function may be obtained for relatively late start times, $a_i/a_f \approx 10$, where a_i and a_f are the initial and final expansion factors.

As can be seen from equation (3), in the limit of asymptotically high initial redshift 1LPT and 2LPT become equivalent since $D_2(a_i)/D_2(a_f) \ll D_1(a_i)/D_1(a_f)$. This has led some to speculate that provided one takes the initial start redshift to be sufficiently high, then it should not matter whether one uses 1LPT or 2LPT. This issue will be investigated in detail in Section 5.

Several earlier studies have explored the importance of 1LPT versus 2LPT initial conditions: Knebe et al. (2009) used GADGET-2 to show that start redshift and 2LPT versus 1LPT had little effect on internal halo properties, specifically testing halo concentration, spin parameter, tri-axiality. They also found little dependence on halo mass or the halo mass function. However, their results may understate any numerical issues because they focused on smaller haloes of $10^{10} - 10^{13} h^{-1} M_\odot$ where the mass function is not very steep, and their statistics were limited due to using low-resolution $N = 256^3$ particles. A more recent study by Jenkins (2010) has shown that there is a definite, although weak, dependence of the subhalo mass function inside Milky Way mass haloes on the choice of the initial conditions.

One last issue, concerning the generation of initial Gaussian random density fields, is that some researches have advocated the use of the 'Hann filter'⁵ (Bertschinger 2001). This is an anti-aliasing filter (Press et al. 1992), and corresponds to setting the Fourier density modes that are outside the Nyquist sphere of the simulation, $k_{\text{Ny}} = \pi N^{1/3}/L$, to vanish by multiplying the transfer function by $W(k) = \cos(\pi k/2k_{\text{Ny}})$ for $k < k_{\text{Ny}}$ and 0 for $k > k_{\text{Ny}}$. The purpose of this is to mitigate some of the anisotropies in the forces due to the cubical lattice. In Section 5 we shall also investigate how the presence of such filtering impacts our goal of an accurate mass function.

Note that for some of the simulations where we test for parameter convergence, we will also make use of a modified version of GRAFIC-2 (Bertschinger 2001). These two initial condition codes were verified to show identical convergence trends with start redshift.

2.2 N-body codes

Once we have obtained an initial condition, we then need to integrate the equations of motion, equation (2). In this study we shall make use of two standard N -body techniques: PKDGRAV V2.2.12 and GADGET-2.

PKDGRAV is our primary simulation code for this study, an early version of which is described in Stadel (2001). The version of the code we use has been MPI parallelized, and uses the hierarchical tree data structure to organize the individual simulation particles. The gravitational force on each particle is calculated using a multipole expansion with Ewald summation to replicate the simulation cube as an approximation of an infinite periodic universe. The peculiar potential around any given particle is obtained from a hexadecapole expansion of the forces. PKDGRAV uses a variable time-step criterion that is synchronized for global time-steps. Particle orbits are integrated with the symplectic leapfrog integrator.

GADGET-2 is a tree-particle-mesh (Tree-PM) code, and full details of which may be found in Springel (2005). The main difference with PKDGRAV is that on large scales it uses Fourier-based methods to solve for the forces and only uses the tree algorithm to solve for forces on small scales. The solution for the potential is then obtained through an interpolation of the PM and tree forces over the force matching region, and typically this is $\sim 4-5$ mesh cells.

In Section 5, we investigate the mass functions from these different codes and explore the convergence properties with different 1LPT and 2LPT start redshifts. Additionally, we aim to determine the typical values for 'generic' run parameters that are required for per cent level convergence. In what follows we shall describe parameters that are mainly specific to PKDGRAV, but will make reference as to how they apply to GADGET-2 or other codes. GADGET-2 parameters are tested further in Smith et al. (2012). The run parameters that we focus on are as follows.

Tree-opening angle Θ . The tree-opening angle controls the accuracy of medium and long range forces. It does this by setting the minimum distance between a given particle and tree node below which the tree node will be 'opened'. Thus the force calculations for a given particle will include contributions from entire nodes and/or individual particles. A discussion of how Θ relates to other tree types can be found in Stadel (2001).

Softening ϵ . In order to avoid excessively large accelerations and hence excessively short time-steps, the small-scale gravitational interactions must be 'softened'. This makes sense for simulations of collisionless systems like CDM, where the particles represent large coarse-grained elements of the microscopic phase space. PKDGRAV and GADGET-2 both use a softened kernel: gravitational forces approach zero for spatially coinciding particles, and become Newtonian at 2ϵ for PKDGRAV and 2.8ϵ for GADGET-2. PKDGRAV uses the K_3 softening kernel of Dehnen (2001) while GADGET-2 uses a spline kernel. The force softening leads to a minimum resolved spatial scale. Throughout, we make use of constant comoving softening.

Time-step η . Each particle is on an adaptive time-step with length proportional to the time-step parameter η . The actual time-step length for each particle is based on the magnitude of its current acceleration $|\mathbf{a}|$, the softening length ϵ , and the time-step parameter η , in accordance with the relation:

$$\Delta t \geq \eta \sqrt{(\epsilon/|\mathbf{a}|)}. \quad (6)$$

This technique allows significant computational savings in cosmological simulations when only a small fraction of particles are in dense regions requiring the shortest time-steps.

⁴ <http://cosmo.nyu.edu/roman/2LPT>

⁵ The 'Hann filter' is sometimes (erroneously, according to Wikipedia) referred to as the 'Hanning' filter.

Table 1. Suite of cosmological simulations. Col. 1: N -body code used. Col. 2: initial condition method, note that we used the 2LPT code throughout, except where there is a -g, which denotes the use of GRAFIC-2; -HF denotes use of a Hann filter. Col. 3: box size. Col. 4: particle mass. Col. 5: initial condition start redshifts that have been simulated. Col. 6: final redshift. Col. 7: time-stepping parameter. Col. 8: force softening, ϵ , in units of the mean inter-particle spacing l_m . Col. 9: tree-opening angle (ERRTOLTHETA for GADGET-2 runs). The additional GADGET-2 parameters were set to: ERRTOLINACCURACY=0.02, MAXRMSDISPLACEMENTFAC=0.2, MAXSIZETIMESTEP=0.02, MINSIZETIMESTEP=0.000, ERRTOLFORCEACC=0.005, TREEDOMAINUPDATEFREQUENCY=0.05, PMGRID=1024.

Code	IC method	L (h^{-1} Mpc)	m_p ($h^{-1} M_\odot$)	z_i	z_f	η	ϵ	Θ
PKDGRAV	1LPT	17.625	3.88×10^5	30, 49, 100, 200, 400	9.8	0.15	$l_m/50$	0.55
PKDGRAV	2LPT	17.625	3.88×10^5	30, 49, 100, 200, 400	9.8	0.15	$l_m/50$	0.55
GADGET-2	1LPT	17.625	3.88×10^5	30, 49, 100, 200, 400	9.8	0.2	$l_m/30$	0.5
GADGET-2	2LPT	17.625	3.88×10^5	30, 49, 100, 200, 400	9.8	0.2	$l_m/30$	0.5
PKDGRAV	2LPT	17.625	3.88×10^5	400	9.8	0.15	$l_m/50$	0.4
PKDGRAV	2LPT	17.625	3.88×10^5	400	9.8	0.15	$l_m/50$	0.68
PKDGRAV	2LPT	17.625	3.88×10^5	400	9.8	0.15	$l_m/50$	0.8
PKDGRAV	2LPT	17.625	3.88×10^5	400	9.8	0.07	$l_m/50$	0.55
PKDGRAV	2LPT	17.625	3.88×10^5	400	9.8	0.2	$l_m/50$	0.55
PKDGRAV	2LPT	17.625	3.88×10^5	400	9.8	0.25	$l_m/50$	0.55
PKDGRAV	2LPT	17.625	3.88×10^5	400	9.8	0.3	$l_m/50$	0.55
PKDGRAV	2LPT	17.625	3.88×10^5	400	9.8	0.6	$l_m/50$	0.55
PKDGRAV	1LPT-g	17.625	3.88×10^5	400	9.8	0.15	$l_m/50$	0.55
PKDGRAV	1LPT-g-HF	17.625	3.88×10^5	400	9.8	0.15	$l_m/50$	0.55
PKDGRAV	1LPT-g	17.625	3.88×10^5	400	9.8	0.15	$l_m/20$	0.55
PKDGRAV	1LPT-g	17.625	3.88×10^5	400	9.8	0.15	$l_m/10$	0.55
PKDGRAV	1LPT-g	17.625	3.88×10^5	400	9.8	0.15	$l_m/5$	0.55
PKDGRAV	1LPT-g	17.625	3.88×10^5	400	9.8	0.15	$l_m/2$	0.55
PKDGRAV	1LPT	2048	6.05×10^{11}	30, 49, 200	0.0	0.15	$l_m/50$	0.55($z > 2$), 0.7($z < 2$)
PKDGRAV	2LPT	2048	6.05×10^{11}	30, 49, 100, 200, 400	0.0	0.15	$l_m/50$	0.55($z > 2$), 0.7($z < 2$)
PKDGRAV	1LPT-g	2048	6.05×10^{11}	400	0.0	0.15	$l_m/50$	0.55($z > 2$), 0.7($z < 2$)
GADGET-2	2LPT	2048	6.05×10^{11}	30, 200	0.0	0.2	$l_m/20$	0.5

In summary, we shall investigate how the halo mass function varies with: the initial start redshift; with 1LPT or 2LPT initial conditions; with Nyquist filtering; we shall explore results for two simulation codes; and how variations in Θ , ϵ , η affect our results. Besides these, we shall also explore finite volume effects and mass resolution.

3 SIMULATIONS

3.1 Simulation suite

We have generated a suite of N -body simulations that are designed to explore the accuracy with which we may estimate the halo mass function and its dependence on how we simulate the dark matter (as discussed in Section 2). All of the simulations that we have performed evolve $N = 1024^3$ equal-mass dark matter particles. We consider periodic cubes of size $L = 17.625 h^{-1}$ Mpc evolved to $z = 10$, and cubes of size $L = 2048 h^{-1}$ Mpc evolved to $z = 0$. The relative box sizes were chosen so that haloes corresponding to $\sim 3\sigma$ fluctuations in the density field are sampled by $N_h \sim 1000$ particles at the final output. This corresponds to $M \sim 3.8 \times 10^8 h^{-1} M_\odot$ for the small boxes at $z = 10$, and $M \sim 6.1 \times 10^{14} h^{-1} M_\odot$ at $z = 0$ for the larger boxes. Thus, the final haloes in the small box simulations are in an evolutionary state similar to the clusters in the large box simulations at $z = 0$.

Although the $L = 17.625 h^{-1}$ Mpc box is very small, the effects of finite volume on our study are attenuated because we examine haloes early, at $z = 10$, when the typical halo mass-scale is still much smaller than the total simulation mass. There is no need to

apply a finite volume correction as in e.g. Reed et al. (2007) to these simulations because each of our convergence series utilizes identical initial conditions and particle displacements. Finite volume effects, to the extent that they can be accounted for with a simple linear correction technique, are thus identical within each convergence series.

The cosmological parameters that we adopted for the small box runs were consistent with 5-year *Wilkinson Microwave Anisotropy Probe* (WMAP5) (Komatsu et al. 2009): $\Omega_m = 0.274$, $\Omega_\Lambda = 0.726$, $\Omega_b = 0.046$, $h = 0.705$, $n_s = 0.96$, $\sigma_8 = 0.812$, where these parameters are the density parameters in matter, vacuum energy and baryons; the dimensionless Hubble parameter; the primordial power spectral index; and the variance of the density fluctuations on scales of $R = 8 h^{-1}$ Mpc. The transfer function that we used to create the initial conditions was produced using the prescription of Eisenstein & Hu (1998). For the large box runs, the cosmological parameters we adopted were consistent with WMAP7 (Komatsu et al. 2011): $\Omega_m = 0.2726$; $\Omega_\Lambda = 0.7274$, $\Omega_b = 0.046$, $h = 0.704$, $n_s = 0.963$, $\sigma_8 = 0.809$. The transfer function for these runs was computed using CAMB (Lewis, Challinor & Lasenby 2000). Full details for the entire suite of simulations are given in Table 1.

3.2 Halo identification

We identify dark matter haloes using the *friends-of-friends* (FoF) algorithm (Davis et al. 1985). This links together all particles that are separated by less than the linking length b , where b is expressed in units of the mean inter-particle separation. Throughout, we use $b = 0.2$, and we use the particular implementation of the algorithm

internal to PKDGRAV; a similar implementation is provided by the code FOF.⁶ The estimated iso-density contour that the FoF recovers is $\delta = \delta\rho/\bar{\rho} = n_c b^{-3} - 1 \approx 81.62$ (More et al. 2011).

In the literature there is a wide range of alternate approaches to the identification of dark matter haloes: e.g. the *spherical over-density* (SO) algorithm (Lacey & Cole 1994; Tinker et al. 2008); 6D phase space algorithms (Behroozi, Wechsler & Wu 2013); and for a review of methods see Knebe et al. (2011). Rather than explore all of these different methods here, we shall work under the assumption that: an accurate FoF mass function implies an accurate mass function for all other algorithms designed to select *approximately* virialized objects. Anecdotal support for this is provided by McBride et al. (2011), who found similar convergence behaviour with simulation set-up for the FoF ($b = 0.2$) and SO 200 mass functions. We shall reserve the task of establishing the veracity of this assumption for future work.

A number of systematic errors have been noted for haloes obtained with the FoF algorithm. First, the recovered halo masses are systematically overestimated with respect to the ‘true’ halo mass (Warren et al. 2006; Lukić et al. 2009; Trenti et al. 2010; More et al. 2011). This owes to the fact that the true halo mass distribution is sampled by a finite number of particles. This mass overestimation has been quantified for spherical mock haloes by Lukić et al. (2009) and more recently by More et al. (2011). Secondly, FoF haloes also experience ‘bridging’, which may occur when two distinct haloes undergo a close encounter. This systematic effect: links unvirialized systems, it acts to reduce the overall number of haloes found, and it predominantly occurs for the highest mass objects and is stronger at higher redshifts (Lukić et al. 2009).

Warren et al. (2006) and Bhattacharya et al. (2011) have provided empirical corrections, determined from cosmological simulations, for the systematic FoF errors. However, these empirical models are mass and redshift independent, which may make them insufficient for our goal of achieving a ~ 1 per cent accurate mass function. We would expect the FoF errors to include dependencies on the specific distribution of mass within haloes, which depends on both mass and redshift. Nevertheless, we assert that the FoF mass overestimation should not affect our convergence tests because they are all performed at the same mass resolution. For these reasons, we use the raw FoF masses, and remark that this will affect our ability to recover an ‘unbiased’, FoF mass function. However, this is sufficient to our purposes of quantifying relative differences in the mass function of different simulations.

4 MASS FUNCTION CONVERGENCE I: SIMULATION PARAMETERS

In this section, we show convergence of the FoF halo mass function with varying simulation run and set-up parameters. We estimate the number of haloes per logarithmic mass interval, $dn(M)/d\log_{10}M$, using a Gaussian kernel in log mass. The Gaussian kernel is convenient for diagnostics because it avoids the ‘saw-tooth’ pattern that emerges in a binned mass function as a result of the discretization of halo masses, particularly at low masses where the simulation particle mass is a significant fraction of the mass-width of a bin. The number of haloes with mass M_k is estimated by the sum:

$$N_k = \frac{\sum_i M_{hi} f_g}{\sum_i M_{hi}}, \quad (7)$$

where f_g is a Gaussian kernel (in $\log_{10}M$) of width $h = 0.0625$, chosen to minimize Poisson fluctuations without introducing systematic error to the mass function. To minimize computational cost, we truncate the kernel beyond a range of $\pm a = 3h$. The number of haloes per logarithmic mass interval per unit volume is:

$$\frac{dn(M)}{d\log_{10}M} = \frac{N_k}{V h \operatorname{erf}(a/\sqrt{2})} \quad (8)$$

where V is the simulation volume. The kernel mass M_k is estimated by the following:

$$M_k = \frac{\sum_i N_{hi} M_{hi} f_g}{\sum_i N_{hi}} \quad (9)$$

analogous to using the average halo mass in a bin.

Poisson errors can be estimated from the Gaussian kernel halo numbers N_{hi} , through use of the formula (Heinrich 2003; utilized for the halo mass function in Lukić et al. 2007):

$$\sigma_{\pm} = \sqrt{N_h + \frac{1}{4} \pm \frac{1}{2}}. \quad (10)$$

In what follows, we will show results for all haloes with more than eight particles per halo. This is done for purely diagnostic purposes, since the poorly resolved haloes are strongly affected by particle shot-noise errors.

4.1 Tree-opening angle Θ

The top panel of Fig. 1 presents the results from our convergence study of the tree-opening angle parameter Θ , used in the code PKDGRAV. These results are for the case of the $z = 10$, $L = 17.625 h^{-1} \text{Mpc}$ runs. The figure clearly shows that, for haloes with masses $M \leq 10^9 h^{-1} M_{\odot}$, the runs with $\Theta \leq 0.68$ are converged at the sub-per cent level. For haloes with masses $M \gtrsim 10^9 h^{-1} M_{\odot}$ ($N_p \gtrsim 3000$), the figure shows that scatter in the mass function begins to dominate our results. This implies that systematic errors at greater than the per cent level cannot be ruled out for more massive haloes, which may be most affected by tree-related criteria. In this case, the $\Theta = 0.8$ run deviates modestly from the other runs. Larger values of Θ have been shown to cause significant direct force errors (Stadel 2001).

The increased scatter in the mass function for haloes with $M \geq 10^9 h^{-1} M_{\odot}$ may seem somewhat surprising, given that all of our simulations had the same initial realization of the density field so that there is no ‘sample variance’ between them. However, even the most accurate particle simulation is still essentially a Monte Carlo representation of a mass distribution. This means that the mass of each halo has a significant uncertainty, which can at best be equal to the square root of the number of its particles. The scatter in the mass function arises from a convolution of the true halo number counts with the scatter associated with simulating, sampling and measuring the masses of haloes (see discussion in Section 3.2). Hence it is non-trivial to determine the true uncertainty. Fig. 1 shows the expected Poisson errors (long thin error bars), and one can see that differences between well-converged runs are at sub-Poisson levels.

For a better estimate of the uncertainty, we create randomly subsampled 1-in-8 particle volumes from the full simulation snapshot and measure the 1σ scatter between the FoF mass functions of the subsampled volumes. In Fig. 1, the results of this exercise are denoted by the short, thick, red error bars. This scatter tends to be an overestimate of the true uncertainty, since the scatter in the FoF mass will be larger in the randomly sampled volume due to the smaller numbers of particles per halo. For the most massive haloes,

⁶ www-hpcc.astro.washington.edu/tools/fof.html

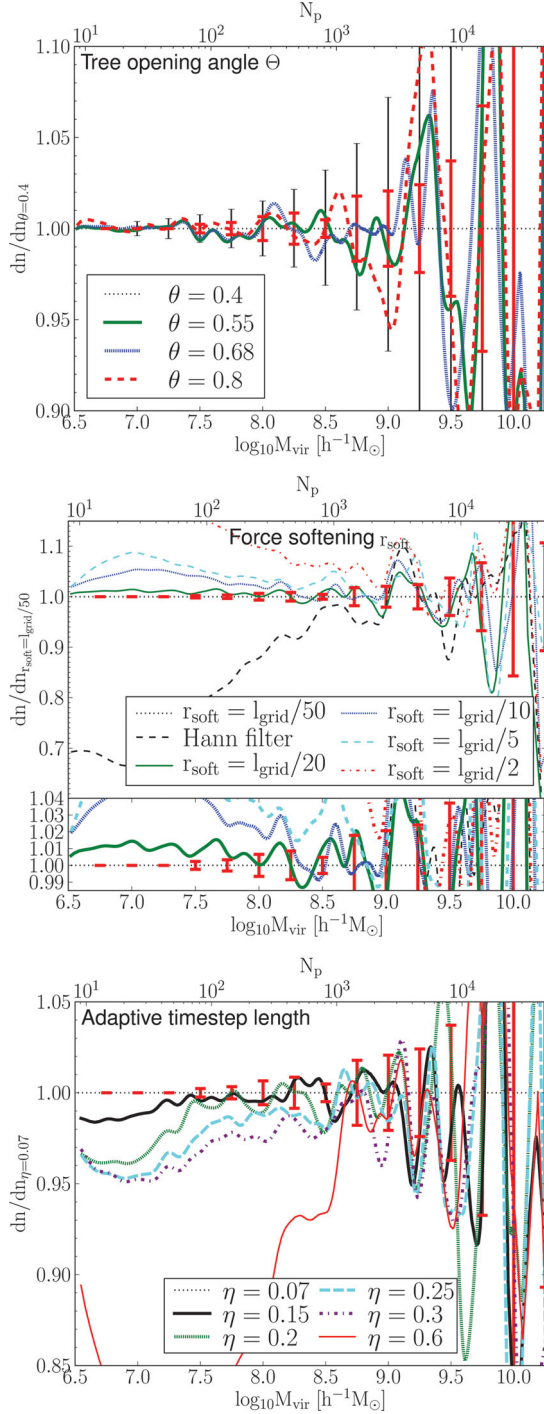


Figure 1. Variation of the halo mass function with simulation parameters, relative to the mass function obtained from our fiducial simulation, as a function of FoF halo mass. All panels show the results from the $L = 17.625 h^{-1} \text{Mpc}$, PKDGRAV simulations at $z = 10$. All runs used the same initial realization of the density field. Long thin error bars show the Poisson errors estimated from the number of haloes. Short thick red error bars are estimated from the scatter of the mass functions of sub-sampled versions of the simulations. The top, middle and bottom panels show the results of variations in tree-opening angle Θ , force-softening ϵ , and adaptive time-step parameter η , respectively. The black dashed line in the middle panel shows the large (and assumedly undesirable) impact of application of the Hann anti-aliasing filter to the initial density field, on the halo mass function. Per cent level convergence is apparent for each run parameter for haloes larger than ~ 100 particles.

this overestimation may become worse due to the increased effects of halo bridging (or unbridging), which has a large effect on halo masses and thus on the inferred mass function. Taking these issues into account, we estimate that we are sensitive to per cent level systematic shifts in the FoF mass function for haloes with less than ~ 3000 particles.

4.2 Force-softening ϵ

The central panel of Fig. 1 presents the results from our convergence study of the force-softening parameter ϵ . As can be seen, the commonly used softening value of $\epsilon = l_m/20$ is converged at near the per cent level over all masses ($l_m = L/N^{1/3}$). We also see that the low-mass end of the mass function is very sensitive to the choice of ϵ . Haloes with $N \gtrsim 1000$ particles appear only weakly dependent on ϵ , provided $\epsilon \lesssim l_m/10$ and within our statistical limitations.

For the cases where $\epsilon \gtrsim l_m/10$, forces do not become Newtonian until beyond the FoF linking length of $l_m/5$. Interestingly, for these large softening lengths, we find that the halo abundances at a fixed mass are increased. One possible explanation for this effect is that the lower central densities of the heavily softened halo profiles (e.g. Moore et al. 1998; Fukushige & Makino 2001; Power et al. 2003; Reed et al. 2005; Tinker et al. 2008) lead to higher densities near the outer edges of haloes. This increased outer-shell density implies larger FoF masses as more particles are linked to the outer layers (Bhattacharya et al. 2011; see Section 6.2). For the smallest haloes with few particles, the virial radii and softening are of similar order. The minimum resolved halo mass thus increases as softening increases (e.g. Lukić et al. 2007). These issues have implications for a common running mode for cosmological simulations where computational speed-up at high redshifts is achieved by using a ‘physical softening’, wherein the comoving softening length is scaled by the expansion factor, with a typical maximum of $f_{\text{softmax}} \sim 10\epsilon$. Effectively, in this mode, the softening is frozen in physical coordinates at scale factor $a = 1/f_{\text{softmax}}$. Our tests suggest that allowing a comoving softening larger than $l_m/20$ at high redshifts leads to a significant numerical error for the early forming haloes, which is likely to affect high-density regions at late times.

Finally, the black dashed line in the central panel of Fig. 1 shows the results of applying the Hann anti-aliasing filter. Whilst it may help to correct errors due to the an-isotropic lattice structure, it introduces a ~ 30 per cent suppression in the number of lower-mass haloes relative to the unfiltered initial conditions run and relative to the expected nearly power-law behaviour of the mass function predicted from theory (Bond et al. 1991). Hann filtered and unfiltered runs only agree at the per cent level for haloes with $N \gtrsim 3000$ particles.

4.3 Time-step size η

The bottom panel of Fig. 1 presents our study of how variations in the adaptive time-stepping parameter η affects the estimated mass functions. The results demonstrate that, for haloes with $M \lesssim 10^9 h^{-1} M_\odot$ (3000 particles), an increase in η leads to a decrease in the abundance of haloes. We find that per cent level convergence in the mass function can be achieved with $\eta \approx 0.15$ for all halo masses, or $\eta \approx 0.2$ for a 100 particle minimum mass. This value is consistent with the value of $\eta = 0.2$ found by Power et al. (2003), who examined the convergence behaviour of the density profiles of dark matter haloes. This similar converged time-step size is not surprising given that low-redshift halo centres consist of some of

the earliest material to be bound into haloes (Diemand, Madau & Moore 2005). For haloes with $N \gtrsim 1000$ particles, the mass function converges with a much longer time-step corresponding to $\eta = 0.6$.

Finally, we note that the value $\eta = 0.2$ for PKDGRAV is similar to the default size of the adaptive time-stepping in GADGET-2: the parameter $\text{ERRTOLINTACCURACY} = \eta^2/2$ has a default setting of 0.025, which corresponds to $\eta = 0.22$.

5 MASS FUNCTION CONVERGENCE II: INITIAL CONDITIONS

5.1 Results: small boxes at $z = 10$

Fig. 2 compares the behaviour of the 1LPT and 2LPT initial conditions, as a function of the start redshift z_i , for the small-box simulations at $z = 10$ evolved with PKDGRAV. The top and middle panels show the ratios of the halo mass functions for different start redshifts with the halo mass function obtained from the start redshift $z_i = 800$, for 1LPT and 2LPT, respectively. The bottom panel presents the ratio of the 1LPT and 2LPT mass functions for simulations with the same start redshift. We see that both the 1LPT and 2LPT initial conditions converge to yield the same mass function as start redshift increases. However, the convergence properties of the 1LPT runs are very slow, whereas the 2LPT runs appear to converge much faster.

For the case of 2LPT, we notice that per cent level convergence can be achieved for haloes with at least ~ 1000 particles and in simulations that have undergone $a_f/a_i \gtrsim 10$ expansions. For 1LPT, the $a_f/a_i = 80$ run ($z_i = 800$) is barely converged down to $N \sim 1000$. For haloes, with $N \lesssim 1000$ particles, even the 2LPT mass function

is poorly converged with z_i for all expansion factors tested. The abundances of small haloes appear to diminish as start redshift is increased; this is apparent in both the 2LPT and 1LPT tests. This suggests that $N \sim 1000$ particles represent a minimum halo mass for stability to z_i .

A curious coincidental feature of the 1LPT initial condition series is that small haloes appear to be converged at the ~ 2 per cent level by $z_i = 200$ (except for the largest masses) and nearly at the ~ 1 per cent level by $z_i = 400$ (except for the smallest masses). The more accurate 2LPT start redshift series confirms that this 1LPT convergence is an illusion. With later start redshift, the increased number of haloes due to more accuracy in the 1LPT initial conditions is offset by independent errors that act to decrease the number of haloes, resulting in *false convergence*. This highlights the fact that convergence is a necessary but not sufficient condition to guarantee accuracy. Some previous studies that appeared to show good mass function convergence with early enough 1LPT initialization, such as Reed et al. (2003) and Reed et al. (2007) (fig. A1), among others, likely also suffered from this false convergence. Our larger particle numbers and corresponding better statistical uncertainty, combined with 2LPT comparisons, allow us to make more robust conclusions.

5.2 Results: large boxes at $z = 0$

In order to confirm that the convergence behaviour of the small box simulations at $z = 10$ can be applied to the cluster regime at lower redshifts, we present the results from tests run to $z = 0$ in the $L = 2 h^{-1}$ Gpc boxes.

Fig. 3 shows the results from the 1LPT and 2LPT z_i convergence simulations run with the code PKDGRAV. As for the case of the

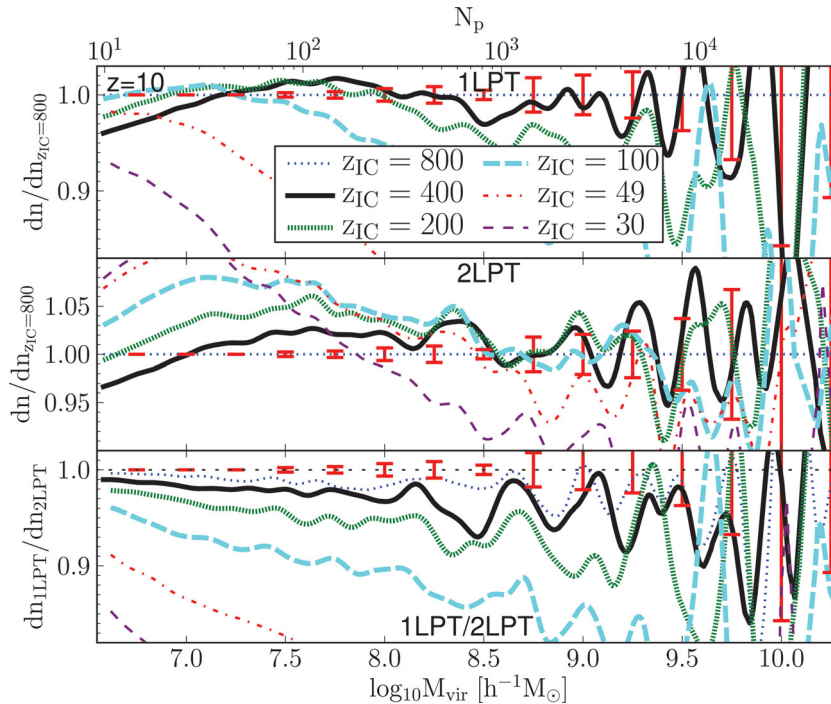


Figure 2. Variation in the halo mass function as a function of start redshift, for 1LPT and 2LPT initial conditions using the N -body code PKDGRAV. Top panel: results for 1LPT initial conditions. Middle panel: results for 2LPT initial conditions. Bottom panel: the ratio of the 1LPT and 2LPT mass functions approaches unity at high initial start redshifts. All panels show results of the $L = 17.625 h^{-1}$ Mpc box at $z = 10.0$. The 1LPT series displays ‘false convergence’ for ~ 100 particle haloes. Per cent level convergence is met for 2LPT for $N_p \gtrsim 1000$ particles to $z_i = 400$ or $a_f/a_i \approx 40$.

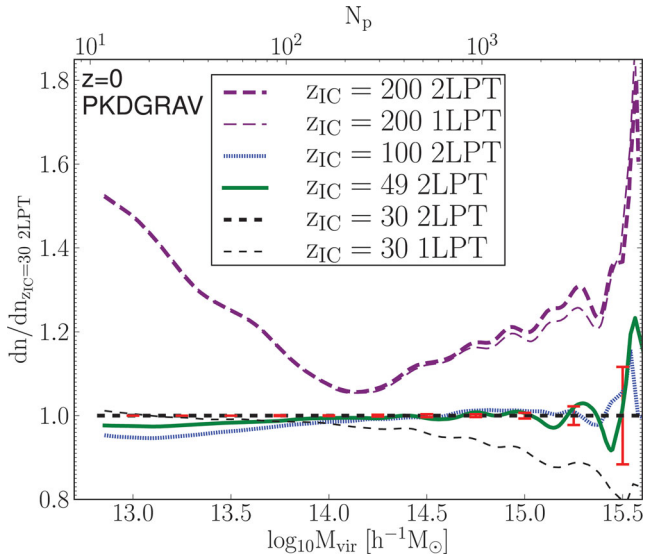


Figure 3. Mass functions from the simulations started with 1LPT and 2LPT initial conditions in the $L = 2h^{-1}$ Gpc boxes at $z = 0$, relative to the simulation started at $z_i = 30$ with 2LPT initial conditions. All simulations here were run with PKDGRAV. Per cent level convergence with 2LPT is found between $z_i = 30$ and $z_i = 49$ runs at ~ 1000 particles. Extremely early starts ($a_f/a_i \gtrsim 100$) lead to serious errors.

small-box simulations at $z = 10$, we see that low-mass haloes are missing with high start redshifts $z \gtrsim 49$; the $z = 0$ ‘pivot’ mass-scale, below which convergence is poor, is somewhat smaller at $N \sim 300$ particles. For larger haloes, the 2LPT mass function is well converged so long as $z_i \lesssim 49$; the $z_i = 100$ curve deviates from the $z_i = 30$ curve at just over the per cent level at this mass scale, so is marginally statistically consistent at the per cent level. A striking feature here is that when a high enough start redshift is used that Zel’Dovich and 2LPT initial conditions are converged, $z_i = 200$, serious errors are present in the $z = 0$ mass function. A likely explanation is that with such a high z_i , cosmological perturbation amplitudes become smaller than the effective amplitude of spurious numerical perturbations. In this $z_i = 200$ run, spurious haloes begin to form at very early times, initially dominated by eight particle structures; visual inspection reveals that the spurious haloes are aligned with the initial grid of particles. The effects of these early spurious haloes lead to the overabundance of haloes at $z = 0$. This underscores the point that 2LPT initial conditions are preferable because they allow one to start at lower redshift where numerical errors are more controllable.

We note that pure PM codes may perform better with high-redshift starts because the PM technique is well suited to following low-amplitude linear perturbations. A tree code (and also a tree-PM code), however, is subject to force errors that may accumulate over time, even if they are small, because these errors are correlated with the tree structure. Tree code force errors could thus seed spurious structures that dominate over real cosmological perturbations when start redshift is very high (and initial cosmological perturbations are very low). Further, the accumulated errors would be expected to worsen if time-step length is decreased. The PM code, although it may perform better at early times, is limited in spatial resolution by the mesh size, which is typically much larger than the softening scale of a tree (or tree-PM) code, making it non-ideal for modelling the internal properties of haloes.

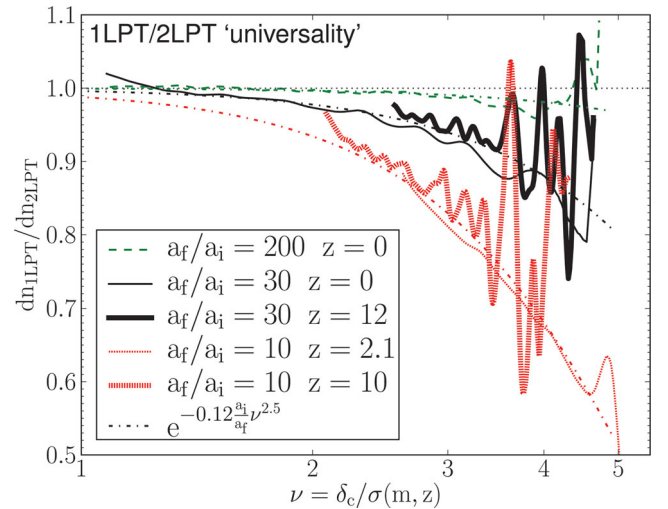


Figure 4. Ratio of 1LPT mass functions with those obtained from 2LPT initial conditions as a function of the equivalent halo peak height ν . The thick solid black and thick hatched red lines show results from the $L = 17.625 h^{-1}$ Mpc boxes at $z \geq 10$, and the thin green dashed, solid black and hatched red show results from the $L = 2 h^{-1}$ Gpc boxes. All results here are for PKDGRAV. The ratio of 1LPT/2LPT mass functions is ‘universal’ in that it depends mainly upon a_f/a_i and ν , independent of the specific halo mass range and redshift, and fit by the dot-dashed line.

5.3 Transformation to universality

It has been noted that when halo masses are translated into equivalent ‘peak-height’, i.e. where $M \rightarrow \nu(M, a)$, then the mass function takes on a ‘universal form’ (Sheth & Tormen 1999). The peak height is defined through the relation $\nu = \delta_c/\sigma(M, z)$, where $\delta_c = 1.686$ is the present-day linearly extrapolated overdensity threshold for collapse in the spherical collapse model, and where $\sigma(M, z)$ is the variance of matter fluctuations on mass scale $M = 4\pi R^3 \bar{\rho}/3$. Note also that owing to the fact that $\sigma(M, a) \propto D(a)$, $\nu \propto D^{-1}(a)$, where $D(a)$ is the linear theory growth factor.

In Fig. 4 we present the ratio of the 1LPT with the 2LPT mass functions for the small box at $z = 10$ and the big box at $z = 0$ as a function of ν . We see that the 1LPT mass function error appears to be relatively independent of mass and redshift for equal values of ν . This universal behaviour is expected in Press–Schechter formalism wherein mean halo formation time depends only upon ν , and δ_c is independent of redshift. We present a fit to the ratio of the 1LPT to 2LPT mass function for our range of data $1 \lesssim \nu \lesssim 5$,

$$\frac{dn_{1LPT}}{dn_{2LPT}} = e^{-0.12 \frac{a_i}{a_f}} \nu^{2.5}. \quad (11)$$

Per cent level convergence between 1LPT and 2LPT initial conditions at scales relevant for the cluster mass function is not achieved until extremely early starts at $a_f/a_i \sim 200$. However, such early start redshifts lead to very small initial density perturbations, which through the relative increase of numerical errors preclude the codes that we tested from being able to accurately follow the growth of structure.

5.4 Comparison of PKDGRAV with GADGET-2

The quest for obtaining mass function predictions accurate to within one per cent requires different N -body simulation codes to provide

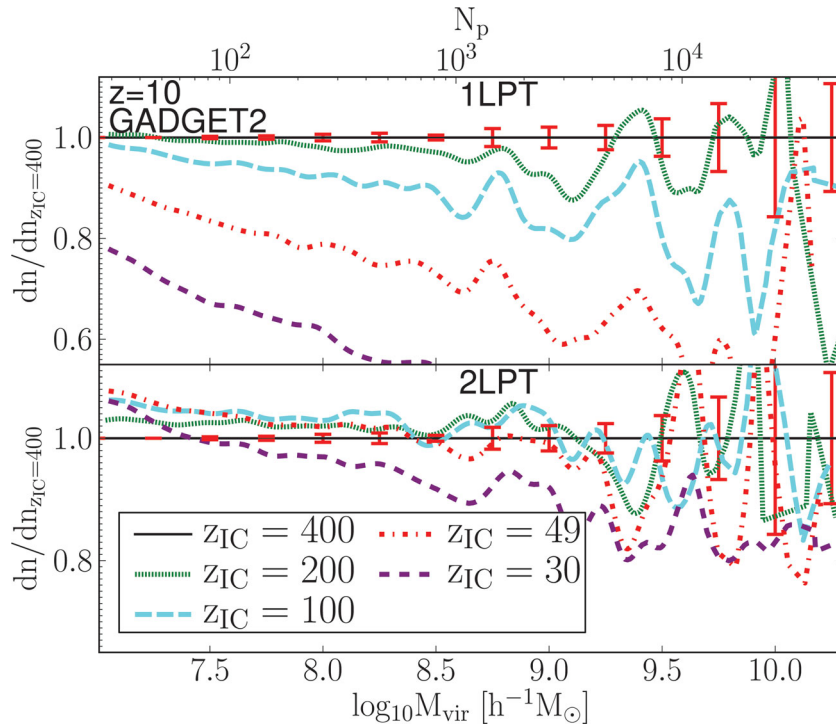


Figure 5. Variation in the halo mass function as a function of start redshift, for 1LPT and 2LPT initial conditions using the N -body code GADGET-2. Top panel: results for simulations started with 1LPT initial conditions. Bottom panel: results for simulations started with 2LPT initial conditions. All panels are for the $L = 17.625 h^{-1}$ Mpc box at $z = 10.0$. Convergence of GADGET-2 runs with initial conditions are very similar to that found for PKDGRAV (Fig. 2).

consistent results at this level. Of course, if results disagree at $\gtrsim 1$ per cent for any two codes, then one would need at least three independent N -body codes to break the degeneracy and so decide which results were correct. Having said that, we now compare the initial condition convergence series of simulations obtained using PKDGRAV with results obtained from the widely used N -body code GADGET-2. Note that we have made no attempt to find ‘optimal’ parameters for GADGET-2, but instead we have adopted some rather generic choices for these. The full list of simulations that we have performed with GADGET-2, including the exact choices for run parameters, is presented in Table 1.

In Fig. 5 we compare the 1LPT and 2LPT initial conditions as a function of z_i , but this time using GADGET-2, plotted here down to the limit $N = 20$ particles. We find that the results exhibit almost identical behaviour to those obtained from the PKDGRAV runs. The small difference is that the suppression of the mass function at low masses with increasing start redshift, apparent in 2LPT runs for haloes with fewer than $N \sim 1000$ particles, is slightly milder with GADGET-2. This appears to enhance the effect of ‘false convergence’ of the 1LPT GADGET-2 simulations.

Fig. 6 shows the ratio of the mass functions obtained from GADGET-2 with those obtained from PKDGRAV. Note that we used identical initial conditions in all cases. We find that GADGET-2 systematically produces up to a 10 per cent higher mass function for low-mass haloes ($N \lesssim 200$) than PKDGRAV. This excess abundance slightly increases with increasing start redshift. We note that the differing mass functions could be a result of differing halo structure, which could lead to systematic differences in FoF masses, and does not necessarily mean the codes are truly producing different numbers of haloes. We also note that GADGET-2 appears to have several per cent fewer haloes at $N \sim 1000$ particles, for the highest start redshifts.

Fig. 7 compares 2LPT mass functions from GADGET-2 and PKDGRAV in the large box at $z = 0$. This figure shows that when the lower redshift start is used, $a_t/a_i = 30$, the two codes generally agree within 2 per cent. However, when high-redshift starts are used, $a_t/a_i \sim 200$, the codes diverge from each other and from the true answer – recall Fig. 3 where we showed that the lower redshift start is converged in PKDGRAV runs.

This code comparison shows that there is a weak systematic shift with start redshift between the codes. However, the convergence behaviour of GADGET-2 with start redshift and of 1LPT versus 2LPT still provides useful verification of the PKDGRAV initial condition tests. Ultimately, a robust comparison of absolute accuracy between codes would require that run parameters for each code are run at self-converged values. The code differences are consistent with the level of agreement found between these same codes in Heitmann et al. (2008) when considering our improved statistics and resolution. Further investigation is warranted to reveal whether the differences between the two codes is caused by inherent differences between the TreePM and the pure tree method or whether they are instead due to the use of non-ideal run parameters in GADGET-2. This is beyond the scope of this paper and we shall reserve a wider study for future work.

6 CONVERGENCE OF OTHER PROPERTIES

In this section we consider the sensitivity of other dark matter statistics: to the adopted simulation parameters; to whether we employ 1LPT or 2LPT initial conditions; and to the adopted start redshift. We shall restrict our exploration to the mass power spectrum and the one-point probability distribution function (PDF) of dark matter density as additional diagnostics for determining simulation accuracy.

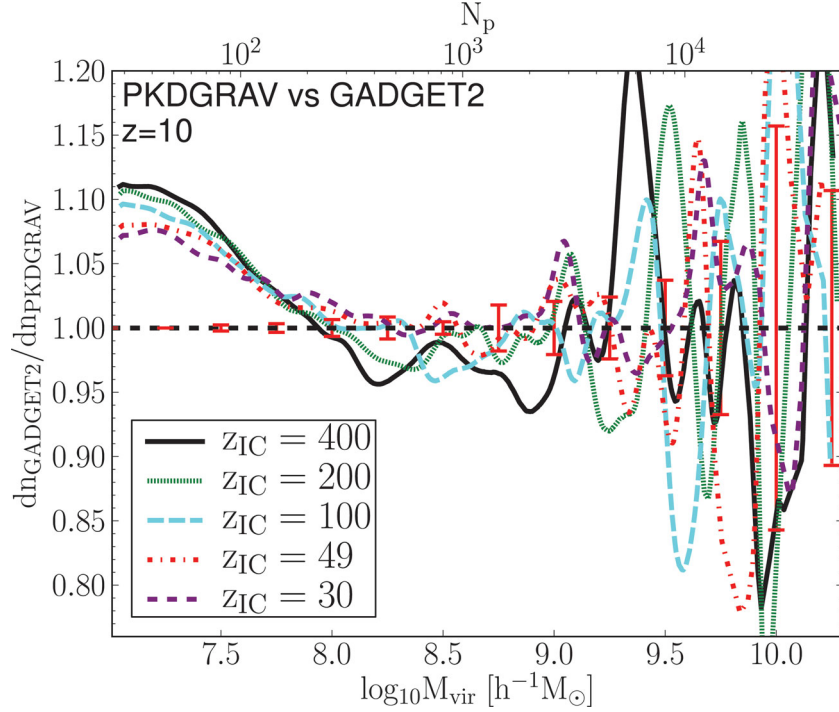


Figure 6. $z = 10$. 2LPT. Mass functions from PKDGRAV and compared with GADGET-2. Some differences are present at small scales, but the relative offset is independent of start redshift, providing some confirmation of our initial condition convergence criteria. We have verified the PKDGRAV, but not the GADGET-2 run parameters for per cent level self-convergence.

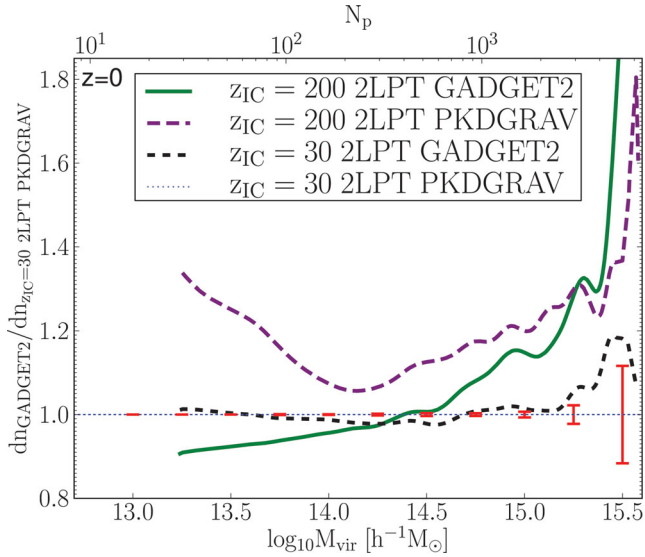


Figure 7. Mass functions from GADGET-2 compared with PKDGRAV in the $L = 2 h^{-1}$ Gpc box. The agreement with GADGET-2 using relatively standard run parameters is at the couple per cent level with $z_i = 30$. Extremely early starts ($a_f/a_i \gtrsim 100$) lead to serious errors independent of the simulation code used.

6.1 Mass power spectra

For a finite cubical patch of the Universe, the matter power spectrum is defined to be:

$$\langle \delta_{k_1} \delta_{k_2}^* \rangle \equiv P(|k_1|) \delta_{k_1, k_2}^K / V_\mu \quad (12)$$

where V_μ is the volume of the patch and δ_k is the discrete Fourier series expansion of the density field. For equal-mass dark matter

particles, the discrete representation of the Fourier space density field can be written as (Peebles 1980):

$$\delta(\mathbf{x}) = \frac{V_\mu}{N} \sum_i^N \delta^D(\mathbf{x} - \mathbf{x}_i) - 1 \Leftrightarrow \quad (13)$$

$$\delta(\mathbf{k}) = \frac{1}{V_\mu} \int d^3\mathbf{x} \delta(\mathbf{x}) e^{i\mathbf{k} \cdot \mathbf{x}},$$

where N is the number of particles. The matter power spectra were estimated for each simulation using the standard Fourier based methods (Smith et al. 2003; Jing 2005; Smith, Sheth & Scoccimarro 2008): particles and halo centres were interpolated on to a 1024^3 cubical mesh, using the Cloud-in-Cell (CIC) algorithm (Hockney & Eastwood 1988); the Fast Fourier Transform of the discrete mesh was computed using the FFTW libraries; the power in each Fourier mode was estimated and then corrected for the CIC charge assignment; these estimates were then bin averaged in spherical shells of logarithmic thickness.

Before we proceed to the results, we note that it is not necessarily the case that a simulation that yields a $\lesssim 1$ per cent accurate halo mass function should also yield a $\lesssim 1$ per cent accurate matter power spectrum, and *vice versa*. Different requirements for simulation parameters are possible because the mass range in our mass functions only involves a small fraction of the total mass in the simulation. And because our estimates of the measured power spectra do not extend to scales as small as the virial radius of the smallest haloes considered.

6.1.1 Variation with simulation parameters

Fig. 8 shows the dependence of the matter power spectrum on the simulation run parameters: (top panel) tree-opening angle Θ ;

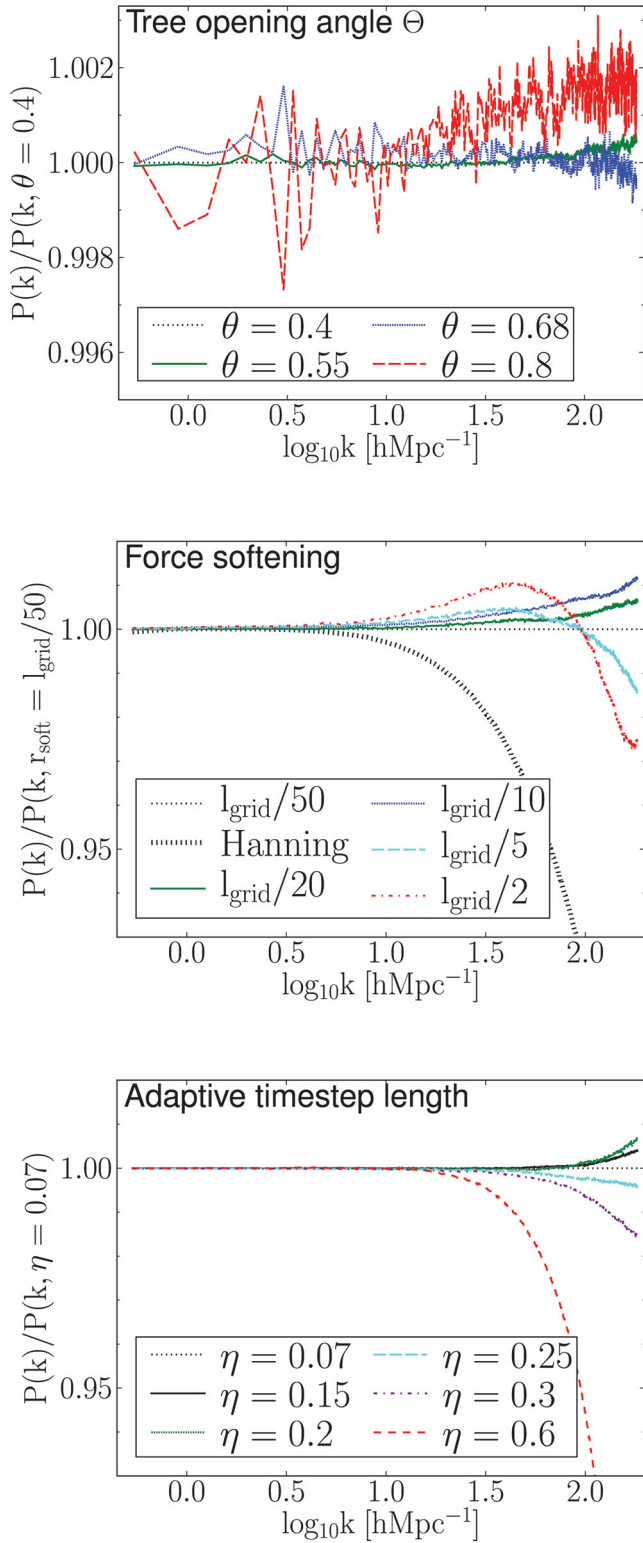


Figure 8. Relative variation of the dark matter power spectra with the simulation run parameters for the N -body code PKDGRAV. Top panel: variation with respect to the tree-opening angle θ . Middle panel: variation with respect to the force-softening parameter ϵ . Bottom panel: variation with respect to the time-step parameter η . All panels show the $L = 17.625$ box at $z = 10$. Per cent level convergence is seen for each run parameter.

(middle panel) the force-softening parameter ϵ ; (bottom panel) time-step parameter η , for the N -body simulation code PKDGRAV. These results show that the estimated power spectra, on large scales⁷ ($k/k_{\text{fun}} < 10$), are only weakly sensitive to variations in the choice of (θ, ϵ, η) . However, on smaller scales, the power spectra show significant deviations. For the force-softening tests, we find that for $\epsilon \leq l_m/5$ the spectra appear to be converged at the sub-per cent level on large scales, with a ‘bump’ at $k/k_{\text{fun}} \approx 100$ and a steep drop at smaller scales. This small-scale drop in power is consistent with the puffing up of halo cores that appears to affect the mass function when softening is large. Similarly, for the case of the time-stepping parameter η , we see that for $\eta = 0.6$ there is a $\gtrsim 1$ per cent suppression of power for $k/k_{\text{fun}} \gtrsim 100$. This can be attributed to the large time-step not being able to follow the rapid changes in the acceleration of particle orbits in the cores of haloes – and hence the failure to capture the complex orbit of particles in dense environments. This discussion is limited to PKDGRAV run parameters; a detailed study of the dependence of the power spectrum on GADGET-2 run parameters can be found in Smith et al. (2012).

6.1.2 Variation with initial conditions: small boxes

Fig. 9 shows the variation of the mass power spectra with the choice of 2LPT or 1LPT initial conditions, and with the adopted initial start redshift for the small box simulations at $z = 10$. The top panel of Fig. 9 shows the results for the 1LPT initial conditions. We observe that the power spectra are only converged on the largest scales. On smaller scales, we find that the power increases with increasing start redshift and that the results are almost converged at the per cent level only after $a_i/a_i \gtrsim 80$ expansions.

The middle panel of Fig. 9 shows the results for the 2LPT initial conditions. Here we find that simulations that were started with $z_i \gtrsim 100$ are converged for $k/k_{\text{fun}} \lesssim 90$. For simulations that possess lower start redshifts we find again a suppression of power, although the effect is much reduced when compared to the equivalent 1LPT start redshift. On smaller scales, $k/k_{\text{fun}} \gtrsim 100$, we find that the power is 1 per cent converged for $100 \leq z_i \leq 400$ ($10 \lesssim a_i/a_i \lesssim 40$), while the $z = 800$ start has up to 2 per cent less power.

The bottom panel of Fig. 9 presents the ratio of the power spectra obtained from the 1LPT simulations with the 2LPT power spectra for various start redshifts. We see that the convergence of the results from the 1LPT simulations with the 2LPT simulations is very slow, and that per cent level convergence is only obtained for $z_i \gtrsim 800$.

Before continuing, we note that, whilst it appears that per cent level convergence in the power spectra may be achieved between 1LPT and 2LPT for very high start redshifts, we have already shown in Section 5 that such high start redshifts are too early to produce an accurate mass function, owing to numerical noise. We are therefore cautious about such convergence.

6.1.3 Variation with initial conditions: large boxes

We now repeat the same set of tests as done for the previous subsection; only this time we now consider the $L = 2048 h^{-1}$ Mpc simulation cubes at $z = 0$.

Fig. 10, bottom panel, shows that the 1LPT and 2LPT initial matter power spectra, measured at $z = z_i$, are converged at the 1 per cent level with respect to each other for the same start redshift.

⁷ Note that owing to the fact that we are comparing results from the $L = 17.5 h^{-1}$ Mpc at $z = 10$ and $L = 2048 h^{-1}$ Mpc boxes at $z = 0$, we shall refer to wavenumbers in units of the fundamental frequency $k_{\text{fun}} = 2\pi/L$.

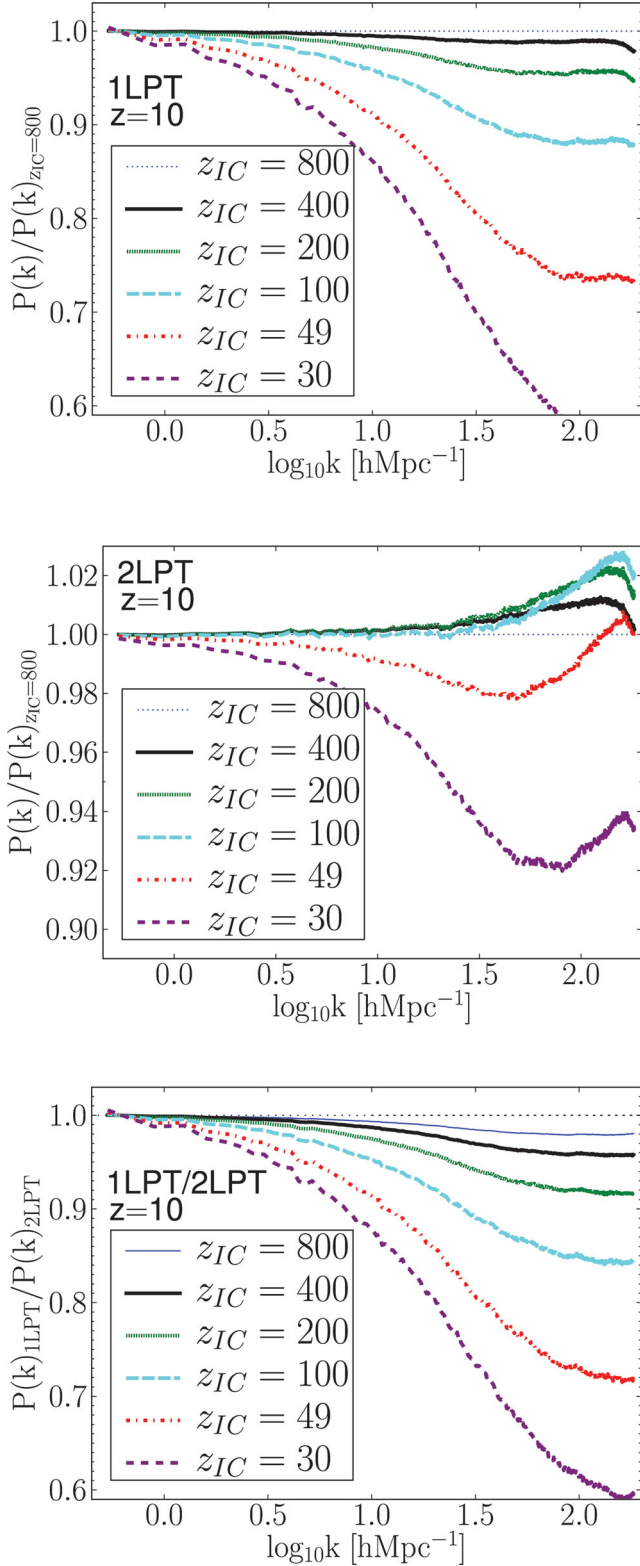


Figure 9. Relative variation of the dark matter power spectra with the initial conditions for the N -body code PKDGRAV. Top panel: variation of spectra with initial start redshift for 1LPT initial conditions. Middle panel: same as the top panel except for the case of the 2LPT initial conditions. Bottom panel: ratio of the power spectra from the 1LPT initial condition runs with respect to the 2LPT runs. All results shown are for the $L = 17.625$ box at $z = 10$. Per cent level convergence is seen for 2LPT.

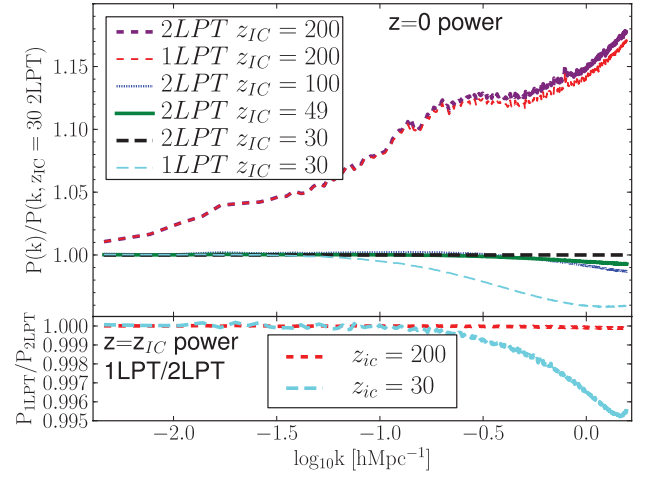


Figure 10. Top panel: ratio of the evolved $z = 0$ matter power spectra for runs with various initial start redshift z_i . The evolved 1LPT and 2LPT power spectra from the early $z_i = 200$ start are very similar to each other, but they lie well above the converged results of lower redshift 2LPT starts. Bottom panel: initial ratio of the 1LPT to 2LPT matter power spectra at two selected initial start redshifts. Differences in 1LPT and 2LPT power are very small (< 1 per cent) at z_i , then grow much larger as the simulation evolves. The simulations of this figure are large box $L = 2048 h^{-1}$ Mpc simulations run with PKDGRAV.

However, as indicated in the top panel of Fig. 10, the evolved spectra started with 1LPT are not converged. On the other hand, the simulations started with the 2LPT initial conditions appear to be converged at the 1 per cent level for $a_f/a_i \leq 100$, except perhaps at the smallest scales, even for start redshifts as low as $z_i \sim 30$. We note that the evolved 1LPT versus 2LPT simulations, started with $z_i = 200$, are significantly discrepant with respect to the other results. As for the case of the mass function, we conjecture that this start redshift is too early for the code PKDGRAV to produce an accurate integration of the equations of motion. This reinforces our earlier findings that 1LPT initial conditions are inadequate for accurate simulations.

Several earlier studies have investigated the importance of 1LPT/2LPT initial conditions on the matter power spectrum (Croce et al. 2006; Heitmann et al. 2010). Our findings are broadly consistent with these studies. However, Heitmann et al. (2010) advocated that 1LPT initial conditions started from $z_i = 200$ would lead to better than 1 per cent precision matter power spectra. Clearly, such a statement is code and run parameter dependent, and one should be careful of increased numerical errors that may allow consistency between 1LPT and 2LPT while still resulting in inaccurate power spectra.

6.2 One-point PDF of matter fluctuations

We now investigate the impact of simulation parameters and 1LPT versus 2LPT initial conditions on the one-point PDF. At high densities, the one-point PDF is a useful probe of the central regions of dark matter haloes, reflecting many properties of a ‘stacked’ halo density profile (e.g. Scherrer & Bertschinger 1991). There are some technical subtleties as to how one computes the one-point PDF, since it requires an estimator for the matter density at a given point. The procedure of estimation in general requires one to smooth the particle distribution and hence the results depend up on adopted smoothing scale (see e.g. Watts & Taylor 2001). Here we have

chosen to compute the one-point PDF with a 64 particle nearest neighbour kernel. This operates in a similar way to the SPH-kernel and constitutes an adaptive smoothing scale.

Fig. 11, top panel, shows the variation of the one-point PDF with the tree-opening angle parameter Θ . We note that the most significant changes are in the regions of highest density, though sensitivity to Θ , beyond the statistical fluctuations, is relatively low. Fig. 11, middle panel, shows the variation of the one-point PDF with the force-softening ϵ . This clearly shows that the effect of too large force softening is to damp the density distribution in the highest density inner regions of dark matter haloes. This ‘puffs up’ haloes, which may explain the increased abundances of lower mass haloes with increased softening length. We also note that as $\epsilon = l/50$ the dense regions appear to be again suppressed. This we attribute to violent two-body encounters that can evaporate halo cores. Fig. 11, bottom panel, shows the variation of the one-point PDF with the time-stepping parameter η . We see that the results are well converged provided $\eta \lesssim 0.15$. The suppression of the high density PDF for large η reinforces our earlier speculation that if η is too large, then the particle orbits in the cores of haloes cannot be integrated sufficiently accurately, damping the densities in the inner regions of dark matter haloes.

Fig. 12 presents the one-point PDF for 1LPT and 2LPT initial conditions for various start redshifts. As can be clearly seen, the results for the 1LPT initial conditions converge very slowly with start redshift. We also note that both the high- and low-density regions appear to be less dense for the simulations that were started with low z_i . For the 2LPT simulations, convergence is reached at much lower start redshifts, roughly $a_f/a_i = 10$ expansion factors of the cube (i.e. around $z_i \sim 100$).

We note all of the converged parameter values are broadly consistent with those that we identified for the mass function in Section 4, though variations in the PDF at the highest densities are generally larger than 1 per cent.

7 DISCUSSION: REMAINING CHALLENGES FOR <1 PER CENT ACCURATE MASS FUNCTIONS

In this section we discuss the remaining challenges that we will have to face in order to approach better than 1 per cent accurate dark matter halo mass functions.

7.1 Mass resolution

In the suite of tests above, we have seen indirect evidence that haloes with fewer than $N \sim 1000$ particles are unlikely to be useful for deriving high-accuracy estimates of the mass function. This suggests that there is a critical mass resolution, below which systematic numerical errors are difficult to control. Interestingly, this resolution is somewhat worse than the $N \sim 300$ particle resolution limit expected for a pure PM code with mesh spacing equal to the initial inter-particle separation (Lukić et al. 2007). It is, however, still better than the more conservative value of $N \sim 2000$ particles proposed by Bhattacharya et al. (2011). Additional support for our claim comes from the work of Trenti et al. (2010), who show through mass-resolution tests that on a halo-by-halo basis, halo masses with $N \lesssim 1000$ particles are systematically too low. Our statistical limitations mean we cannot rule out the possibility that haloes resolved with more $N \gtrsim 3000$ particles might be accurate over a larger range in a_f/a_i than what we find for smaller haloes. In a subsequent paper, we will examine directly the mass resolution convergence.

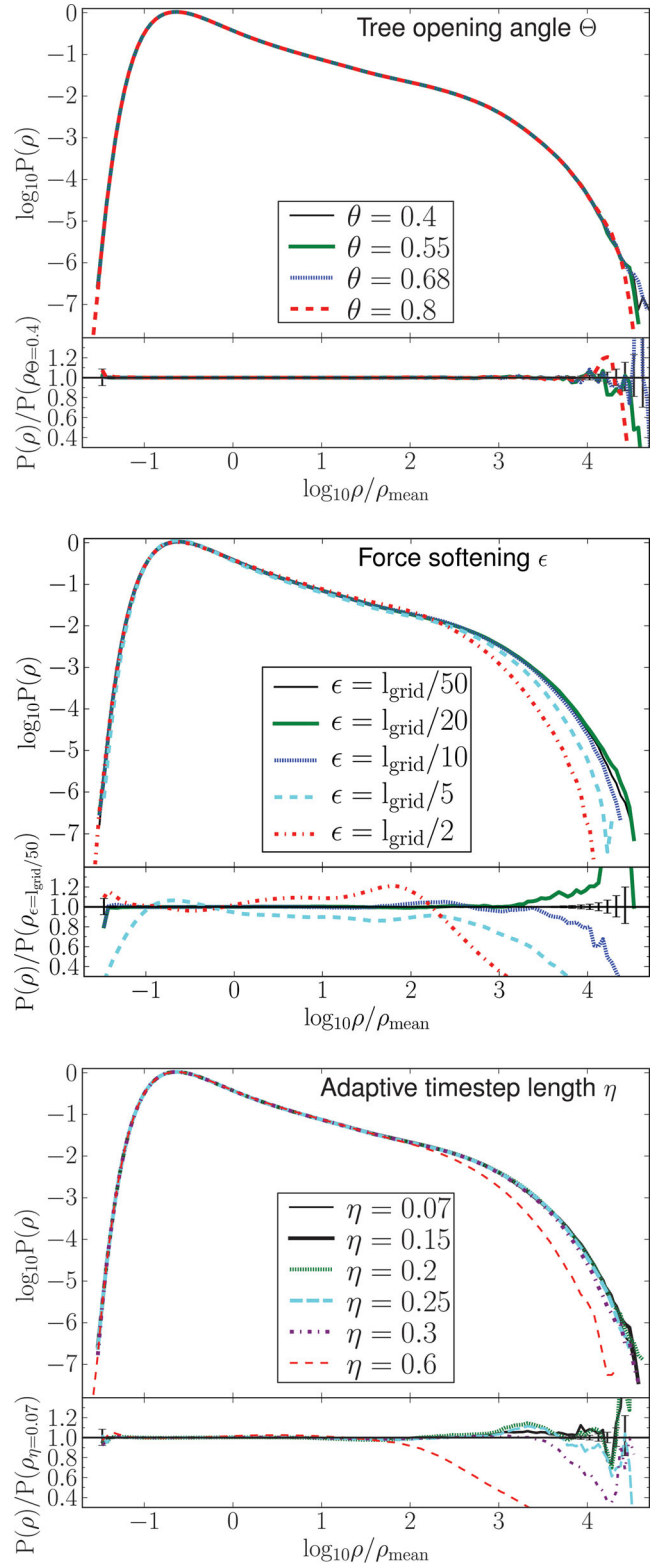


Figure 11. Variation of the one-point PDF of dark matter density fluctuations with the simulation parameters: Top panel, tree-opening angle Θ ; middle panel, force-softening ϵ ; and bottom panel, time-stepping parameter η . All results here were obtained from the small $L = 17.625 h^{-1}$ Mpc simulations at $z = 10$ using the tree-code PKDGRAV. Sensitivity to numerical parameters is greatest at the highest densities.

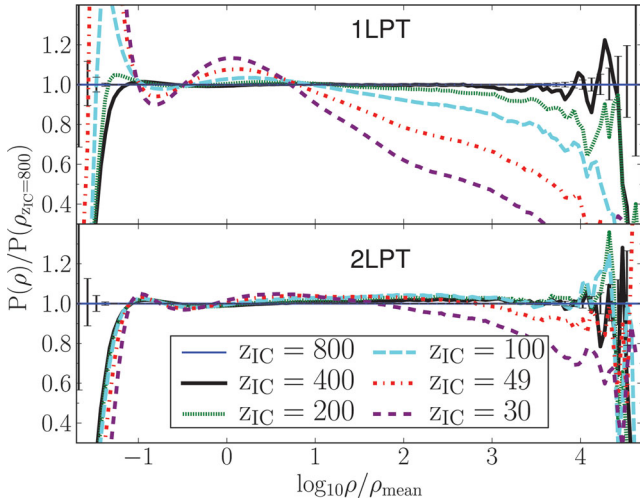


Figure 12. Variation of the one-point PDF of dark matter density fluctuations with respect to the choice of either 1LPT or 2LPT initial conditions for various start redshifts. Top panel: results for the simulations started with 1LPT initial conditions. Bottom panel: the same but for the case of 2LPT initial conditions. These results were obtained from the small $L = 17.625 h^{-1} \text{ Mpc}$ simulations at $z = 10$ using the tree-code PKDGRAV.

If the critical resolution limit of $N \sim 1000$ particles for accurate halo statistics is upheld, then this suggests that the tree code technique may not have much advantage over the PM code technique in recovering an accurate gravity-only mass function. Of course, the higher force resolution for tree codes enables better modelling of higher density regions.

7.2 Statistical precision

The full-sky volume out to $z = 2$ is $V_{\mu} \simeq 200 h^{-3} \text{ Gpc}^3$. This sets the requirement on the minimum simulation volume needed to replicate the survey volume accessible to future cluster surveys. This would correspond to a single simulation cube with a side length of roughly $L \gtrsim 3.67 h^{-1} \text{ Gpc}$, or if one wants to replicate the full sky-light cone, then one would require a cube of length $L \sim 8 h^{-1} \text{ Gpc}$. Obviously performing such a huge simulation with sufficient mass resolution to obtain $N \sim 1000$ particles per halo, for all haloes with masses $M \gtrsim 10^{13.5} h^{-1} M_{\odot}$, would be very challenging prospect. Such total volumes may be cheaply covered by combining the results from many smaller volume simulations. However, individual simulation boxes must be large enough to avoid systematic errors due to mode-discreteness near the box-scale and due to the lack of super-box scale power (a ‘DC-mode’ can help for multiple realization ensembles; Sirko 2005).

An estimate of the minimum box size required to avoid suppressing massive halo formation can be made by computing the effect of missing power at wavelengths larger than the box on the (empirically fit) analytic form of the mass function. Fig. 13 (top panel) demonstrates that a simulation box of roughly $L \sim 2 h^{-1} \text{ Gpc}$ box should be able to capture the mass function at sub-per cent level for all haloes with masses $M \lesssim 10^{16} h^{-1} M_{\odot}$ at $z = 0$. However, in a single realization of such a volume, statistical accuracy will be much lower; Poisson errors remain larger than 1 per cent well below $10^{15} h^{-1} M_{\odot}$. One caveat is that this calculation underestimates the finite volume effects due to the discreteness of modes near the box scale (see e.g. Reed et al. 2007; Smith & Marian 2011).

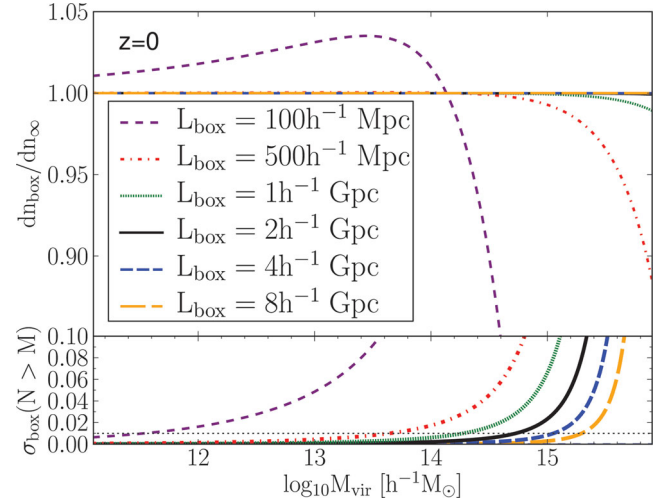


Figure 13. Estimates of the relative dependence of the mass function on simulation cube length compared to the case of an infinite cube. The dependence of the mass function on L is calculated by assuming that the power on scales $k < k_{\text{fun}} = 2\pi/L$ is zero. The bottom panel is the expected Poisson error of the cumulative halo number count within a single simulation volume. Predictions were obtained using the mass function of Bhattacharya et al. (2011).

7.3 Verification of absolute accuracy

The task of verifying that we have actually obtained the true answer is somewhat circular, since it implies that we know already what the true answer is. On our path towards the true answer, we should consider the possibility that our simulations suffer from some level of false convergence. One obvious approach to addressing this issue will be to verify that independent simulation codes give the same results at the desired level of accuracy. However, this does not take into account the pernicious systematic errors, such as the false convergence with redshift that we observed for the 1LPT simulations. In the future, a more complete approach would provide us with a theoretical framework for objectively quantifying ‘accuracy’ and enable us to identify directions in parameter space that would allow us to approach our desired goal.

There are a number of systematic errors that one will need to characterize in detail. One in particular is that associated with the coarse graining of phase space. Spurious perturbations related to mass discreteness may lead to the collapse of small structures around lattice points (Melott 1990; Joyce & Marcos 2007b; Joyce, Marcos & Baertschiger 2009). This effect is well-known for Warm Dark Matter simulations (Wang & White 2007; Schneider et al. 2012), and it therefore must also be present in CDM simulations. Other effects relating to mass discreteness in particle codes have been discussed by a number of authors (e.g. Splinter et al. 1998; Joyce & Marcos 2007a; Romeo et al. 2008).

We have attempted to steer away from making direct statements regarding whether particular papers had errors and by how much because quantifying the accuracy of the works of other authors would require repeating their simulations with exactly the same codes and run parameters. A number of widely used fits in the literature used 1LPT initial conditions, many of them with a lower start redshift than expected to be required for good agreement between 1LPT and 2LPT ICs. In those cases, a systematic underabundance of the halo mass function, especially at high masses or high redshifts, is implied.

An informative comparison can be made from the fact that the fitting formulae of multiple authors are in reasonable agreement with each other. For example, there is agreement at the 2–3 per cent level between the FoF mass function fits of Crocce et al. (2010, using 2LPT ICs), Bhattacharya et al. (2011, using a mix of 2LPT and 1LPT with a high start redshift) and Angulo et al. (2012, using 2LPT), over a wide mass range. Also, there is $\simeq 5$ per cent agreement between the Tinker et al. (2008, 1LPT ICs) and the recent Watson et al. (2012, 1LPT ICs with a higher start redshift) spherical overdensity mass function fits. There is a caveat that such comparisons are only useful to the extent that the different studies do not suffer from the same systematic errors.

If we consider the widely used Tinker et al. (2008) mass function, the authors stated statistical accuracy of 5 per cent is comparable to the systematic error, estimate from equation (11), that we would expect from their results due to their use of 1LPT initial conditions with $a_f/a_i \sim 50$, although this error would approach 10 per cent at the highest masses, while being smaller at lower masses ($\lesssim 10^{15} h^{-1} M_\odot$). They point out some dependence on start redshift between some of their simulations, and exclude from their fit those with the lowest start redshifts due to a systematic underabundance of haloes. We can thereby deduce that the systematic errors in that study due to initial conditions are generally within their quoted statistical accuracy.

7.4 Narrow scale factor range for accuracy

In Section 5, we showed that significant errors are introduced in the halo mass function when the total expansion of the box lies outside the limits ($10 \lesssim a_f/a_i \lesssim 50$). Of course, one would expect that the epoch of a particular halo's formation is more important than final simulation output for assessing whether that halo, and by extension, the mass function, has been modelled accurately. A brief supporting argument is that any errors introduced to early structures are unlikely to 'evolve away', though suppression of the number of early forming haloes may become less noticeable, at fixed mass, after more haloes have formed at lower redshifts. The implication is that haloes in any particular cosmological simulation can be modelled accurately only for a range in formation redshifts of $\Delta_{1+z} \sim 5$. It is possible that this restriction may become less severe as the mass resolution of the simulation is increased. The logic being that more cosmological power at small scales may enable earlier start redshifts to be simulated without leading to an increase in the amount of spurious structure. And, as shown in Section 5, the mass function sensitivity to start redshift is smaller for haloes resolved with more particles. The upper limit of the allowed start redshift range may be code dependent. For example, as we discussed in Section 5.2, a particle-mesh technique may allow higher start redshifts, but typically comes at the cost of worse force resolution.

The narrow range in implied a_f/a_i presents a challenge for simulations with very large dynamic range, wherein it would be difficult to model accurately the mass function of massive cluster haloes forming at $z \sim 0$, while also capturing accurate evolution of the early generations of dwarf galaxy haloes forming already at $z \sim 10$. Even though the fraction of mass assembled into galactic haloes at such early times is small, early galaxy formation occurs preferentially in Lagrangian regions where clusters will later form. These early-forming galactic haloes should be modelled accurately because their feedback processes could have significant effects on the eventual baryon and total cluster mass content through energy injection and preheating, which could begin very early (e.g. Benson & Madau 2003).

7.5 Impact of baryons

We have purposely ignored the important effects of baryons on the halo mass function. Recent hydrodynamic simulations have shown the range of plausible baryon effects on total cluster halo masses to be up to $\simeq 15$ per cent within the radius enclosing 500 times the critical density of the universe (Stanek, Rudd & Evrard 2009). Even for 'adiabatic simulations' wherein gas cooling, star formation and feedback are ignored, baryons may have up to ~ 7 per cent effects on the halo mass function (Cui et al. 2012).

Baryon influences thus present a serious challenge for per cent level accuracy in the mass function required for planned dark energy missions. However, accurate gravity-only simulations, as we have explored in this work, are a pre-requisite for future simulations with more complete baryon physics aiming to obtain per cent level accuracy in the mass function or other properties. To save computational cost, cosmological constraints might rely on a combination of baryon and gravity-only simulations. For example, hydrodynamic simulations of limited cosmological volumes could be used to calibrate the effects of baryons on haloes as well as to derive relations between haloes and observable properties. Then, large volume gravity-only simulations might be utilized to map the dependence of halo numbers and other properties on cosmology.

7.6 Calibrating mass–observable relationship

A further difficulty in using the halo mass function for cosmology is that any practical halo definition one might use, such as through the FoF or SO algorithms, typically have no direct observable counterpart. For example, the emission from X-ray clusters is determined by the (square of) the gas density distribution. One might argue that this leads to more spherical clusters the abundance of which may be better matched to SO haloes in simulations, rather than to FoF haloes (see discussion in Kravtsov & Borgani 2012). Though weak lensing may help to calibrate halo masses (Marian et al. 2009; Mandelbaum et al. 2010; Becker & Kravtsov 2011), optical, X-ray, and Sunyaev–Zel'dovich cluster masses have large scatter and systematic uncertainties (Angulo et al. 2012).

Mock observations from simulations may thus be a superior means of obtaining a more accurate mass function in the observable plane, especially once baryon properties are better modelled. Ultimately, an observationally useful and accurate (cluster) mass function will involve modelling baryon physics and observable properties (or mock observations) and represents a formidable challenge for cosmology – a challenge that must be solved in order to fully exploit future and even current cluster surveys of the Universe.

8 CONCLUSIONS AND GUIDELINES FOR ACCURATE SIMULATIONS

In this paper we have explored the dependence of the mass function of dark matter haloes on simulation run parameters and initial conditions. Our aim has been to perform convergence tests that will illuminate the path to obtaining per cent level accuracy in this statistic. This will be a requirement for future cluster surveys of the Universe that aim to help constrain the nature of dark energy or dark gravity.

In Section 2 we gave a brief overview of the simulation method, paying special attention to how one sets up initial conditions, either using Zel'dovich approximation (1LPT) or 2LPT. We described the simulation codes that we have employed PKDGRAV and GADGET-2, with the former being the main code used throughout this study.

Table 2. *Approximate* run and initial condition parameters that permit per cent level simulation convergence in the halo mass function extracted from gravity-only simulations. 2LPT initial conditions are required. Run parameters Θ , ϵ and η are the tree-opening angle, force softening and time-stepping parameter. η_G denotes the GADGET-2 time-step parameter $\text{ERRTOLINTACCURACY} = \eta^2/2$. a_f/a_i is the ratio of initial to final scale factor at which halo properties are to be considered. L is an estimate of the minimum box length needed to avoid systematic errors in the mass function while V_μ is the comoving volume of the universe accessible to future cluster surveys.

	Θ	ϵ	η	η_G	a_f/a_i	N per halo	L (h^{-1} Gpc)	$V_\mu(z < 2)$ (h^{-3} Gpc ³)
Min. value	–	$l_m/50$	–	–	10	1000	2.0	~ 200
Max. value	0.7	$l_m/20$	0.15–0.20	0.02	40–50	–	–	–

In Section 3 we described the large suite of N -body simulations that we have been performed to study these problems. All simulations were run at $N = 1024^3$ and we covered two regimes: high redshift ($z = 10$), small scale ($L = 17.625 h^{-1}$ Mpc) and low redshift ($z = 0$), large scale ($L = 2 h^{-1}$ Gpc).

In Section 4 we explored the dependence of the mass functions on the simulation parameters and found the following: the resultant mass functions were rather insensitive to the choice of the tree-opening angle, provided $\Theta < 0.7$; the results for haloes resolved with fewer than $N \sim 1000$ particles were sensitive to the choice of force softening, with larger values tending to increase the abundance of haloes in this regime; results were fairly insensitive to the size of the adaptive time-step parameter and that 1 per cent converged results could be achieved for $\eta \lesssim 0.15$. We also demonstrated that the use of anti-aliasing filters, such as the Hann filter, to set up initial conditions, can lead to ~ 30 per cent suppression in the abundance of haloes resolved with $N \lesssim 1000$ particles. We do not advocate the use of the Hann filter, since *there is no aliasing* in the initial conditions to correct.

In Section 5 we performed a detailed study of the impact of the choice of initial conditions on the mass function. We found that the results from simulations that are initialized with 1LPT converge very slowly as the start redshift is increased. The effect of too low a start redshift being the suppression of the formation of high-mass haloes. Furthermore, for the large box simulations, we also found simulations started at very high redshifts $z_i \gtrsim 200$ would fail to correctly follow the build up of structure due to the relative increase in numerical noise. Furthermore, 1LPT initial conditions exhibit ‘false convergence’ with increasing start redshift. Simulations starting from 2LPT initial conditions proved to have very good convergence properties and for simulations that underwent 10–50 expansion factors, yielded per cent level convergence in the halo mass function at the 1024^3 resolution of our tests. We made a direct comparison of these results obtained from integration of the initial conditions with the tree-code PKDGRAV with results from the Tree-PM code GADGET-2, and found almost identical behaviour. However, a detailed comparison of the mass functions from the two codes revealed that GADGET-2 produced an ~ 10 per cent increase in the mass function for haloes resolved with $N \lesssim 10^2$ particles. These results extend and support the earlier findings of Crocce et al. (2006).

In Section 6 we explored the convergence properties of two other statistics of the density field, namely the matter power spectrum and the one-point PDF of matter fluctuations. We found that the simulation parameters that produced $\lesssim 1$ per cent convergence in the mass function would also lead to good convergence behaviour in these statistics. In addition, too high a start redshift for either 1LPT or 2LPT initial conditions would lead to systematic errors. On the other hand, the results from the simulations run with 2LPT initial conditions demonstrated excellent convergence behaviour.

In summary, Table 2 presents a general recipe for the parameters needed for per cent accuracy of the mass function within a gravity-only simulation using a tree code. Except for the tree-opening angle Θ , which has some dependence on the specific tree used, all the other run parameters can be applicable to other tree codes. This list shows required values but is not complete. In future work, one would expect this table to be extended to include the following: if PM forces are used for large-scale force computation, then parameters controlling their accuracy, such as the size of the PM grid, should be included; multipole expansions are used to compute the tree forces, and different codes use different orders: which order is sufficiently accurate for our purposes? Also, there should be some entry associated with the parameters that control the halo finder (halo definition).

Ultimately, inferring cosmological parameters from the cluster mass function will require a number of other issues to be solved relating to baryons and observable properties. Among the difficulties that baryons pose is the gravitational coupling of baryon processes to dark matter (Somogyi & Smith 2010; van Daalen et al. 2011). Inferring observable properties from the simulations for comparison via mass–observable relations or by direct mock catalogues is a further complexity. Thus, per cent level accuracy in numerical simulations represents a formidable challenge, but one that we must meet if future surveys of the Universe are to live up to their potential.

ACKNOWLEDGEMENTS

We thank the anonymous referee for helpful suggestions that have improved this work. We thank Martin Crocce, Cameron McBride, Roman Scoccimarro and Romain Teyssier for helpful discussions. We also thank Salman Habib, Katrin Heitmann and Zarijah Lukić for early discussions that influenced this work. We thank Roman Scoccimarro for making 2LPT public; Volker Springel for making GADGET-2 public, and for providing his B-FOF halo finder; and Ed Bertschinger for making GRAFIC-2 public. The simulations were performed on Rosa at the Swiss National Supercomputing Center (CSCS), and the zbox3 and Schrödinger supercomputers at the University of Zurich. RES acknowledges support from a Marie Curie Reintegration Grant and the Alexander von Humboldt Foundation.

REFERENCES

- Alimi J.-M. et al., 2012, preprint (arXiv:1206.2838)
- Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
- Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
- Bagla J. S., Ray S., 2005, MNRAS, 358, 1076
- Barkana R., Loeb A., 2004, ApJ, 609, 474
- Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25

- Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, *ApJ*, 762, 109
- Benson A. J., Madau P., 2003, *MNRAS*, 344, 835
- Bertschinger E., 1999, *Astrophysics Source Code Library*, Record ascl:9910.004 COSMICS: Cosmological Initial Conditions and Microwave Anisotropy Codes
- Bertschinger E., 2001, *ApJS*, 137, 1
- Bhattacharya S., Heitmann K., White M., Lukić Z., Wagner C., Habib S., 2011, *ApJ*, 732, 122
- Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, *ApJ*, 379, 440
- Bouchet F. R., Colombi S., Hivon E., Juszkiewicz R., 1995, *A&A*, 296, 575
- Buchert T., 1994, *MNRAS*, 267, 811
- Buchert T., Melott A. L., Weiss A. G., 1994, *A&A*, 288, 349
- Carbone C., Fedeli C., Moscardini L., Cimatti A., 2012, *J. Cosmology Astropart. Phys.*, 3, 23
- Courtin J., Raseria Y., Alimi J.-M., Corasaniti P.-S., Boucher V., Füzfa A., 2011, *MNRAS*, 410, 1911
- Crocce M., Pueblas S., Scoccimarro R., 2006, *MNRAS*, 373, 369
- Crocce M., Fosalba P., Castander F. J., Gaztañaga E., 2010, *MNRAS*, 403, 1353
- Cui W., Borgani S., Dolag K., Murante G., Tomatore L., 2012, *MNRAS*, 423, 2279
- Cunha C., Huterer D., Doré O., 2010, *Phys. Rev. D*, 82, 023004
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371
- Dehnen W., 2001, *MNRAS*, 324, 273
- Diemand J., Madau P., Moore B., 2005, *MNRAS*, 364, 367
- Efstathiou G., Davis M., White S. D. M., Frenk C. S., 1985, *ApJS*, 57, 241
- Eisenstein D. J., Hu W., 1998, *ApJ*, 496, 605
- Fukushige T., Makino J., 2001, *ApJ*, 557, 533
- Haiman Z., Mohr J. J., Holder G. P., 2001, *ApJ*, 553, 545
- Heinrich J., 2003, Technical Report, Coverage of Error Bars for Poisson Data CDF note 6438. Batavia, Fermilab
- Heitmann K. et al., 2008, *Comput. Sci. Discovery*, 1, 015003
- Heitmann K., White M., Wagner C., Habib S., Higdon D., 2010, *ApJ*, 715, 104
- Hockney R. W., Eastwood J. W., 1988, *Computer Simulation using Particles*. Hilger, Bristol
- Iliev I. T., Mellema G., Shapiro P. R., Pen U.-L., Mao Y., Koda J., Ahn K., 2012, *MNRAS*, 423, 2222
- Jenkins A., 2010, *MNRAS*, 403, 1859
- Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, *MNRAS*, 321, 372
- Jing Y. P., 2005, *ApJ*, 620, 559
- Joyce M., Marcos B., 2007a, *Phys. Rev. D*, 76, 103505
- Joyce M., Marcos B., 2007b, *Phys. Rev. D*, 75, 063516
- Joyce M., Marcos B., Baertschiger T., 2009, *MNRAS*, 394, 751
- Knebe A., Wagner C., Knollmann S., Dieckhoff T., Krause F., 2009, *ApJ*, 698, 266
- Knebe A., The Haloes Gone Mad Team, 2011, *MNRAS*, 415, 2293
- Komatsu E., The WMAP Team, 2009, *ApJS*, 180, 330
- Komatsu E., Smith K. M., Dunkley J., The WMAP Team, 2011, *ApJS*, 192, 18
- Kravtsov A., Borgani S., 2012, *ARA&A*, 50, 353
- Lacey C., Cole S., 1994, *MNRAS*, 271, 676
- Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
- Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, 538, 473
- Lima M., Hu W., 2004, *Phys. Rev. D*, 70, 043504
- Lima M., Hu W., 2005, *Phys. Rev. D*, 72, 043006
- Lukić Z., Heitmann K., Habib S., Bashinsky S., Ricker P. M., 2007, *ApJ*, 671, 1160
- Lukić Z., Reed D., Habib S., Heitmann K., 2009, *ApJ*, 692, 217
- McBride C. et al., 2011, *BAAS*, 43, 24907
- Mandelbaum R., Seljak U., Baldauf T., Smith R. E., 2010, *MNRAS*, 405, 2078
- Marian L., Bernstein G. M., 2006, *Phys. Rev. D*, 73, 123525
- Marian L., Smith R. E., Bernstein G. M., 2009, *ApJ*, 698, L33
- Marian L., Hilbert S., Smith R. E., Schneider P., Desjacques V., 2011, *ApJ*, 728, L13
- Matarrese S., Verde L., Jimenez R., 2000, *ApJ*, 541, 10
- Melott A. L., 1990, *Comment. Astrophys.*, 15, 1
- Moore B., Governato F., Quinn T., Stadel J., Lake G., 1998, *ApJ*, 499, L5
- More S., Kravtsov A. V., Dalal N., Gottlöber S., 2011, *ApJS*, 195, 4
- Oguri M., Takada M., 2011, *Phys. Rev. D*, 83, 023008
- Peebles P. J. E., 1980, in Treiman S. B., ed., *The Large-scale Structure of the Universe*. Research Supported by the National Science Foundation. Princeton University Press, Princeton, NJ, p. 435
- Pillepich A., Porciani C., Reiprich T. H., 2012, *MNRAS*, 424, 44
- Planck Collaboration, 2011, preprint (arXiv:1101.2026)
- Power C., Knebe A., 2006, *MNRAS*, 370, 691
- Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, *MNRAS*, 338, 14
- Press W. H., Schechter P., 1974, *ApJ*, 187, 425
- Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, *Numerical Recipes in FORTRAN. The Art of Scientific Computing*, 2nd edn. University Press, Cambridge
- Reed D., Gardner J., Quinn T., Stadel J., Fardal M., Lake G., Governato F., 2003, *MNRAS*, 346, 565
- Reed D., Governato F., Verde L., Gardner J., Quinn T., Stadel J., Merritt D., Lake G., 2005, *MNRAS*, 357, 82
- Reed D. S., Bower R., Frenk C. S., Jenkins A., Theuns T., 2007, *MNRAS*, 374, 2
- Romeo A. B., Agertz O., Moore B., Stadel J., 2008, *ApJ*, 686, 1
- Rozo E. et al., 2009, *ApJ*, 699, 768
- Rozo E. et al., 2010, *ApJ*, 708, 645
- Scherrer R. J., Bertschinger E., 1991, *ApJ*, 381, 349
- Schneider A., Smith R. E., Maccio A. V., Moore B., 2012, *MNRAS*, 424, 684
- Scoccimarro R., 1998, *MNRAS*, 299, 1097
- Sehgal N., The ACT Collaboration, 2011, *ApJ*, 732, 44
- Sheth R. K., Tormen G., 1999, *MNRAS*, 308, 119
- Shimon M., Rephaeli Y., Itzhaki N., Dvorkin I., Keating B. G., 2012, *MNRAS*, 427, 828
- Sirko E., 2005, *ApJ*, 634, 728
- Smith R. E., Marian L., 2011, *MNRAS*, 418, 729
- Smith R. E. et al., 2003, *MNRAS*, 341, 1311
- Smith R. E., Sheth R. K., Scoccimarro R., 2008, *Phys. Rev. D*, 78, 023523
- Smith R. E., Reed D. S., Potter D., Marian L., Crocce M., Moore B., 2012, preprint (arXiv:1211.6434)
- Somogyi G., Smith R. E., 2010, *Phys. Rev. D*, 81, 023524
- Splinter R. J., Melott A. L., Shandarin S. F., Suto Y., 1998, *ApJ*, 497, 38
- Springel V., 2005, *MNRAS*, 364, 1105
- Stadel J. G., 2001, PhD thesis, AA (University of Washington)
- Stanek R., Rudd D., Evrard A. E., 2009, *MNRAS*, 394, L11
- Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, *ApJ*, 688, 709
- Trenti M., Smith B. D., Hallman E. J., Skillman S. W., Shull J. M., 2010, *ApJ*, 711, 1198
- van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, *MNRAS*, 415, 3649
- Vanderlinde K., The SPT Collaboration, 2010, *ApJ*, 722, 1180
- Vikhlinin A. et al., 2009, *ApJ*, 692, 1060
- Wang L., Steinhardt P. J., 1998, *ApJ*, 508, 483
- Wang J., White S. D. M., 2007, *MNRAS*, 380, 93
- Wang S., Haiman Z., Hu W., Khoury J., May M., 2005, *Phys. Rev. Lett.*, 95, 011302
- Warren M. S., Abazajian K., Holz D. E., Teodoro L., 2006, *ApJ*, 646, 881
- Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2012, preprint (arXiv:1212.0095)
- Watts P. I. R., Taylor A. N., 2001, *MNRAS*, 320, 139
- Wu H.-Y., Zentner A. R., Wechsler R. H., 2010, *ApJ*, 713, 856
- Zel'Dovich Y. B., 1970, *A&A*, 5, 84